

MÉTODOS DE ANÁLISIS DEL COMPORTAMIENTO DE LA OFERTA
DE TRABAJO A PARTIR DE LA REGRESIÓN CON DATOS
DE PANEL AUTONÓMICOS: UNA METODOLOGÍA
DE ESTIMACIÓN NO PARAMÉTRICA

Xosé Manuel Martínez Filgueira
A CORUÑA 1998

**MÉTODOS DE ANÁLISIS DEL COMPORTAMIENTO DE LA
OFERTA DE TRABAJO A PARTIR DE LA REGRESIÓN
CON DATOS DE PANEL AUTONÓMICOS: UNA
METODOLOGÍA DE ESTIMACIÓN NO PARAMÉTRICA.**

Tesis de doctorado presentada por

Xosé Manuel Martínez Filgueira

bajo la dirección del profesor doctor

D. Carlos M^a Fernández-Jardón Fernández

A meus pais.

AGRADECIMIENTOS

Mi gratitud al Profesor Dr. D. Carlos M^a Fernández-Jardón Fernández, quien me sugirió y dirigió este trabajo, además de apoyarme durante todo el tiempo de su realización.

Al Profesor Dr. D. Alejandro Quintela del Río, por su paciente asesoramiento sobre los temas relacionados con los métodos de estimación no paramétrica.

Un agradecimiento colectivo a todas aquellas personas que de una o de otra manera me ayudaron y me apoyaron en la realización de esta tarea, dado que la relación de todas ellas sería excesivamente larga, e injusta, por lo incompleta de la lista.

AGRADECIMIENTOS

Mi gratitud al Profesor Dr. D. Carlos M^a Fernández-Jardón Fernández, quien me sugirió y dirigió este trabajo, además de apoyarme durante todo el tiempo de su realización.

Al Profesor Dr. D. Alejandro Quintela del Río, por su paciente asesoramiento sobre los temas relacionados con los métodos de estimación no paramétrica.

Un agradecimiento colectivo a todas aquellas personas que de una o de otra manera me ayudaron y me apoyaron en la realización de esta tarea, dado que la relación de todas ellas sería excesivamente larga, y injusta, por lo incompleta de la lista.

MÉTODOS DE ANÁLISIS DEL COMPORTAMIENTO DE LA OFERTA DE TRABAJO A PARTIR DE LA REGRESIÓN CON DATOS DE PANEL AUTONÓMICOS: UNA METODOLOGÍA DE ESTIMACIÓN NO PARAMÉTRICA.

1	Introducción.	1
2	Estimación en datos de panel.	8
2.1	Introducción a los métodos de estimación con datos de panel.	10
2.2	Introducción a la estimación no paramétrica y semiparamétrica en los modelos económicos.	19
2.3	Estimación semiparamétrica con datos de panel.	43
3	La oferta de trabajo en las comunidades autónomas.	52
3.1	Las comunidades autónomas en España.	53
3.2	Definición del modelo de la oferta de trabajo.	61
3.3	Evolución de las variables que caracterizan la oferta de trabajo.	77
3.4	Modelización de las series de la oferta de trabajo en las comunidades autónomas.	94
4	Estimación de la oferta de trabajo en las comunidades autónomas.	101
4.1	Estimación mínimo-cuadrática de la oferta de trabajo.	102
4.2	Estimación semiparamétrica de la oferta de trabajo.	138
5	Conclusiones.	150

Apéndice II: Demostración de las propiedades del estimador semiparamétrico.	157
Apéndice III.1: Procedimientos empleados en la elaboración de los datos.	215
Apéndice III.2: Datos empleados.	219
Apéndice III.3: Gráficas de las variables	220
Apéndice III.4: Contraste de los modelos I(1)	244
Apéndice III.5: Resultados de las series de la oferta de trabajo.	245
Apéndice IV.1: Contraste de especificación de Hausman	246
Apéndice IV.2: Resultados de las estimaciones mínimo-cuadráticas de los modelos.	248
Apéndice IV.3: Diseño del estimador semiparamétrico: Cálculos y diagrama de flujo.	249
Apéndice IV.4: Resultados de las estimaciones semiparamétricas de los modelos.	272
Bibliografía	273

ÍNDICE DE GRÁFICAS

Gráfica 1: Nivel de salarios I.....	77
Gráfica 2: Nivel de salarios II.....	78
Gráfica 3: Nivel de salarios III.....	79
Gráfica 4: Nivel de salarios IV.....	81
Gráfica 5: Nivel de educación.....	82
Gráfica 6: P.I.B. por habitante I.....	85
Gráfica 7: P.I.B. por habitante II.....	86
Gráfica 8: Tasa de actividad I.....	89
Gráfica 9: Tasa de actividad II.....	92
Gráfica 10: Impulsos en las series de la oferta de trabajo.....	94
Gráfica 11: Escalones en las series de la oferta de trabajo.....	95
Gráfica 12: Escalones truncados en las series de la oferta de trabajo.....	97
 Gráficas III.1.1 a III.1.18: Nivel de educación de cada autonomía.	
Andalucía, Asturias, Aragón	220
Baleares, Canarias, Cantabria	221
Castilla – La Mancha, Castilla y León, Cataluña	222
Comunidad Valenciana, Extremadura, Galicia	223
Madrid, Murcia, Navarra	224
País Vasco, La Rioja, España	225

Gráficas III.2.1 a III.2.18: Nivel de salarios de cada autonomía.

Andalucía, Asturias, Aragón	226
Baleares, Canarias, Cantabria	227
Castilla – La Mancha, Castilla y León, Cataluña	228
Comunidad Valenciana, Extremadura, Galicia	229
Madrid, Murcia, Navarra	230
País Vasco, La Rioja, España	231

Gráficas III.3.1 a III.3.18: P.I.B. por habitante de cada autonomía.

Andalucía, Asturias, Aragón	232
Baleares, Canarias, Cantabria	233
Castilla – La Mancha, Castilla y León, Cataluña	234
Comunidad Valenciana, Extremadura, Galicia	235
Madrid, Murcia, Navarra	236
País Vasco, La Rioja, España	237

Gráficas III.4.1 a III.4.18: Tasa de actividad de cada autonomía.

Andalucía, Asturias, Aragón	238
Baleares, Canarias, Cantabria	239
Castilla – La Mancha, Castilla y León, Cataluña	240
Comunidad Valenciana, Extremadura, Galicia	241
Madrid, Murcia, Navarra	242
País Vasco, La Rioja, España	243

ÍNDICE DE CUADROS

Cuadro 1: Variación del nivel de salarios.....	84
Cuadro 2: Variación del nivel de educación.....	84
Cuadro 3: Tasas de actividad promedios de las autonomías en 1980.....	93
Cuadro 4: Tasas de actividad promedios de las autonomías en 1985.....	93
Cuadro 5: Tasas de actividad promedios de las autonomías en 1990.....	93
Cuadro 6: Análisis de las series de la oferta de trabajo I.....	99
Cuadro 7: Análisis de las series de la oferta de trabajo II.....	100
Cuadro 8: Coeficientes de determinación entre las variables españolas y autonómicas.....	112
Cuadro 9: Porcentajes de variación de las variables autonómicas.....	125
Cuadro 10: Estadísticos de la oferta de trabajo.....	127
Cuadro 11: Coeficientes estimados para las variables en el modelo 2.....	128
Cuadro 12: Coeficientes estimados para las variables en el modelo 3.....	129

ÍNDICE DE MAPAS

Mapa 1: Nivel de salarios.....	80
Mapa 2: Nivel de educación.....	83
Mapa 3: P.I.B. por habitante.....	87
Mapa 4: Tasas de actividad promedios en 1980.....	91
Mapa 5: Tasas de actividad promedios en 1985.....	91
Mapa 6: Tasas de actividad promedios en 1990.....	91
Mapa 7: Significación de los coeficientes estimados en el modelo 2: Nivel de educación.....	130
Mapa 8: Significación de los coeficientes estimados en el modelo 2: Oferta de trabajo.....	130
Mapa 9: Significación de los coeficientes estimados en el modelo 2: P.I.B. por habitante.....	130
Mapa 10: Significación de los coeficientes estimados en el modelo 3: Nivel de educación.....	131
Mapa 11: Significación de los coeficientes estimados en el modelo 3: P.I.B. por habitante.....	131

1. INTRODUCCIÓN.

Cuando se observa el comportamiento de una variable en varios individuos o entidades a través del tiempo, la posibilidad de que exista una heterogeneidad propia de los individuos que no sea medida por los regresores hace que sea necesario emplear en estos casos modelos de datos de panel.

La existencia de más y mejores bases de paneles de datos impulsó en los últimos años la investigación en problemas relacionados con ellos. En particular, en España, el desarrollo del estado de las autonomías propició el interés por estudiar las realidades autonómicas de cara a disponer de una información regional normalizada, tanto con vistas a la obtención de fondos estructurales derivados de nuestra pertenencia a la Unión Europea como para tener unos criterios objetivos con los que elaborar una política que permita reducir las desigualdades existentes entre las comunidades. Estas necesidades provocaron el aumento y mejora de los paneles de datos sobre las comunidades autónomas y la necesidad de desarrollar y mejorar los métodos de estimación en datos de panel para superar los problemas que aparecen al trabajar en este nuevo contexto.

Dentro de esta tendencia, se desarrolla en este trabajo un nuevo método semiparamétrico de estimación de la regresión en datos de panel. En dicho método se estiman los parámetros a partir de una suavización previa de los datos, realizada mediante una regresión no paramétrica, en lugar de hacerlo directamente de ellos, como hacen los métodos basados en Mínimos Cuadrados. En general, los métodos no paramétricos permiten suavizar las suposiciones previas sobre el modelo, lo que

hace que el método desarrollado mantenga su validez en condiciones menos restrictivas que los métodos clásicos.

Una utilidad de los paneles de datos es que proporcionan un mayor número de observaciones para estudiar un fenómeno o modelo económico. Así, el Estado de las Autonomías y el interés por las estadísticas de las comunidades que trae aparejado van a proporcionar extensos paneles de datos en los que se reflejarán, con frecuencia, situaciones diferentes. Por ese motivo, la consideración de dichas fuentes de datos empíricos, junto con las metodologías de análisis propuestas, ayudarán a comprobar modelos económicos de una manera más completa.

Siguiendo este enfoque, se hace en esta memoria un estudio del comportamiento de la oferta laboral en las comunidades autónomas. En el cual, además de analizar independientemente su evolución temporal, se construye un modelo con tres variables explicativas, las cuales intentarán aproximar los niveles de salarios, educación y riqueza de cada comunidad. A partir del modelo construido se analizarán las posibles diferencias existentes entre comunidades, las cuales proporcionarán, a su vez, una información que se podrá emplear en el mejor conocimiento del modelo.

La elección de este tema se debe al interés que tiene hoy en día el estudio de los mecanismos del mercado laboral, interés originado por la gran extensión del desempleo, y que está provocando un intenso debate en la sociedad sobre sus causas y sus soluciones.

Así pues, esta tesis conlleva un estudio de las propiedades de un estimador de la regresión con datos de panel basado en técnicas no paramétricas. Posteriormente se realiza un análisis sobre un modelo de oferta laboral en España y las comunidades autónomas, en el que se emplean técnicas de datos de panel para construir un modelo global y determinar el efecto que sobre ella tienen tres variables que aproximan los niveles de salarios, de educación y de riqueza de cada comunidad, contrastando si su comportamiento es homogéneo en todas ellas. La información procedente de estos modelos autonómicos será analizada posteriormente, de cara a conocer otros factores que podrían estar influyendo en el mercado de trabajo y no habían sido considerados previamente en el modelo. Finalmente se realiza una aplicación de las técnicas no paramétricas expuestas en el *capítulo segundo* al estudio de este modelo.

A continuación se detalla brevemente el contenido de cada uno de los capítulos, resaltando la importancia que los resultados obtenidos tienen dentro de su propio marco de referencia.

El *capítulo 2* introduce la teoría de la estimación que se usará en la parte empírica del trabajo: las técnicas de regresión en datos de panel y los métodos no paramétricos y semiparamétricos mediante suavización no paramétrica previa de estimación de la regresión.

La idea de estos métodos no paramétricos es estimar una función de regresión mediante un promedio ponderado de los valores de Y que le corresponden a cada

valor de X . Según la función que empleemos para hacer la ponderación obtendremos diferentes métodos de estimación no paramétricos.

La idea de la estimación semiparamétrica con suavización previa consiste en minimizar la distancia entre los valores del modelo teórico considerado y los valores obtenidos mediante el estimador no paramétrico, en lugar de hacerlo con respecto a los datos observados.

El estimador sería entonces aquel que minimiza algún tipo de distancia entre la estimación no paramétrica y la función paramétrica elegida.

Hasta ahora, los trabajos realizados sobre este método de estimación se han aplicado a modelos de regresión, ya con datos transversales, ya con series de tiempo. Sin embargo no se encontraban trabajos en la literatura sobre suavización no paramétrica con datos de panel. En este sentido, la metodología desarrollada en esta memoria supone una aportación interesante. Además se demuestran sus propiedades de consistencia y se estudia su distribución asintótica.

El tercer capítulo está dedicado a la elaboración de un modelo que se contrastará con los datos empíricos.

Se comienza con la presentación del Estado de las Autonomías y comentarios sobre la situación económica de España y las comunidades durante el periodo estudiado (1980-1990).

En el segundo apartado se describe el modelo de oferta laboral en las comunidades autónomas, y en el se considera la posibilidad de que pueda ser influida por tres

variables:

- El nivel salarial, para el que se estudiará la existencia de dos posibles efectos: el *efecto sustitución*, que aumenta la oferta cuando aumentan los salarios, y el *efecto renta* que actúa en sentido contrario. Existe además la posibilidad de que los salarios no afecten directamente a la oferta, sino que sea esta la que condicione su comportamiento. Se estudia en este trabajo cual de las dos hipótesis se adapta mejor al comportamiento de la economía en las comunidades autónomas;
- El nivel educativo, para el que se analizarán posibles reducciones de la oferta laboral causadas por el aumento del tiempo que los jóvenes dedican a su formación, o posibles incrementos producidos por el mayor coste de oportunidad que para personas preparadas tiene el no estar trabajando;
- El nivel económico o de riqueza de la autonomía, cuyas oscilaciones de esta variable pueden atraer o apartar gente del mercado laboral, al actuar sobre la demanda de trabajo.

Previamente a realizar la estimación de dicho modelo se estudia la evolución de las variables construidas, mencionándose los hechos que más tarde podrán afectar a su estimación. Además se realiza un estudio de raíces unitarias, que en el caso de la oferta de trabajo se combina con un análisis Box-Jenkins, obteniendo los hechos relevantes del comportamiento de esta variable en cada comunidad autónoma, y sus posibles discrepancias.

Los datos disponibles constan de 17 submuestras (una por cada comunidad) de 44 datos cada una (un dato por cada trimestre del periodo estudiado).

El *capítulo 4* está formado por la estimación y contrastación empírica del trabajo.

Se comienza (*apartado 4.1*) con la estimación, para cada autonomía y para España, del modelo elaborado en el capítulo anterior. Para España se realizarán dos estimaciones, una empleando los datos procedentes del conjunto de las autonomías y otra con datos agregados. El objetivo de esta doble estimación es estudiar los cambios que pueden inducirse a partir del uso de paneles de datos para la estimación de un modelo económico, así como sus ventajas e inconvenientes. Por otro lado, comparando estos resultados (autonómicos y españoles) con los obtenidos para el modelo español a partir de datos globales, se podrá analizar el efecto que sobre él tiene la existencia de características diferenciadas en partes importantes del territorio.

Para estudiar la existencia de estas características diferenciadas se contrastará, mediante un test de significación de Hausman, qué comunidades se apartan significativamente del modelo global. Una vez detectado el conjunto de autonomías cuyo comportamiento es menos homogéneo, se prescindiría de sus datos y se obtendría un conjunto de datos de panel con el cual estimar más eficientemente el modelo global.

Finalmente se utilizará el conjunto de todas las estimaciones realizadas para elaborar una descripción completa del modelo de la oferta de trabajo en España y la influencia que sobre ella tienen las tres variables consideradas.

El siguiente apartado está dedicado a la estimación semiparamétrica. Se emplea como estimador no paramétrico previo un estimador kernel n -dimensional y se elige cual es la ventana óptima mediante validación cruzada. Con este método se realiza una nueva estimación de los modelos anteriores, comparando sus resultados con los que se obtienen haciendo la estimación por mínimos cuadrados ordinarios. Esto permitirá comprobar si las hipótesis de normalidad de las perturbaciones ha afectado a la estimación mínimo-cuadrática. Si bien la principal aportación que se pretende hacer con este método estaría en el desarrollo de una alternativa a la estimación mínimo-cuadrática cuando los contrastes habituales indiquen esta posibilidad.

Para finalizar, el *capítulo 5* es un resumen de las conclusiones obtenidas a lo largo de esta memoria.

2. ESTIMACIÓN CON DATOS DE PANEL.

El estudio comparativo de relaciones entre variables en distintas comunidades nos ofrece una matriz de datos temporales y a la vez transversales, conocidos como datos de panel.

En los últimos años se ha producido un considerable incremento en la investigación de problemas teóricos y aplicados relacionados con los datos de panel. Este incremento ha sido debido en parte a la mayor disponibilidad de este tipo de datos existente en la actualidad. Ashenfelter, Solon (1982) y Borus (1982) hacen inventarios sobre bases de datos de panel existentes de interés para los economistas.

A la hora de estudiar las estimaciones correspondientes a estos modelos aparecen discusiones introductorias sobre el tema en la mayoría de los manuales de Econometría como por ejemplo Judge. (1988), Kmenta (1986), Novales (1993); visiones más completas aparecen en Hsiao (1986), Chamberlain (1984), Heckman, Singer (1985), Dielman (1989), Dormont (1989); recopilaciones de artículos sobre el tema aparecen en Mazodier (1978), Klevmarken (1989) Hartog, Ridder, Theeuwes (1990), Raj, Baltagi (1992), Maddala (1992).

En las referencias consultadas no hemos encontrado trabajos que analicen los datos de panel sin hacer suposiciones sobre las leyes de distribución de los errores, lo que es conocido como estimación semiparamétrica (Robinson (1988)). Este enfoque será la principal aportación de este capítulo.

Para ello se dedica la primera parte a introducir los modelos básicos de estimación clásica con datos de panel, introduciendo a continuación los métodos semiparamétricos en el ámbito económico. Hasta ahí se establece la situación en el

momento actual. Finalmente se aporta la metodología semiparamétrica con datos de panel, estableciendo sus principales propiedades.

2.1 INTRODUCCIÓN A LOS MÉTODOS DE ESTIMACIÓN CON DATOS DE PANEL.

Cuando un modelo verifica una serie de propiedades fundamentales que permitan, por lo menos, la consistencia de los estimadores, es conveniente hacer uso del mayor número posible de datos, generalmente mejorarán los errores de la estimación.

Sin embargo, cuando surge la duda de que en los datos existan comportamientos heterogéneos o provengan de distintas poblaciones los métodos de estimación deben adaptarse a esa circunstancia. Es en este ámbito en el que aparece la necesidad de estimar mediante datos de panel.

Los datos de panel surgen cuando observamos el comportamiento de un conjunto de variables aleatorias en M entidades durante T períodos de tiempo y estamos interesados en contrastar si existen o no diferencias en el comportamiento de las relaciones entre las variables para los distintos individuos o momentos en los que fue observada. El interés de contrastar esa homogeneidad se debe a que aceptarla sin más puede conducir a inconsistencia o estimaciones erróneas en parámetros de interés.

Además existen otras ventajas para el empleo de datos de panel, como por ejemplo el hecho de que proporciona al investigador un gran número de observaciones, mejorando así la eficiencia de los estimadores; también permiten hacer análisis de hipótesis económicas complejas controlando las influencias debidas tanto a heterogeneidades individuales como temporales.

Estas y otras ventajas son discutidas en más detalle en Chamberlain (1984), Hsiao (1985,1986), Maddala (1987a,b), Klevmarken (1989), Raj, Baltagi (1992).

2.1.1 El diseño de un modelo económico con datos de panel.

El planteamiento básico que subyace a un modelo de regresión en datos de panel consiste en una variable aleatoria (X, Y) $(p+1)$ -dimensional, en la que se representa por X a la variable exógena p -dimensional y por Y a la variable dependiente. Esta variable es observada en M individuos o entidades durante T períodos de tiempo, por lo que tendremos $M \cdot T$ observaciones, siendo $(X_{1t}, X_{2t}, \dots, X_{pt}, Y_t)$ el valor que toma la variable (X, Y) para el individuo r en el instante t .

Si suponemos que la relación entre Y y X es de tipo lineal:

$$Y = \alpha + X\beta + \varepsilon$$

donde α es un parámetro, $\beta = (\beta_1, \beta_2, \dots, \beta_p)'$ un vector p -dimensional de parámetros y ε una variable aleatoria con media cero. Vamos a tener entonces distintos modelos según existan diferencias o no entre los distintos individuos.

Para simplificar el modelo vamos a asumir que los parámetros son constantes a través del tiempo y que pueden variar a través de los individuos, así podemos considerar un modelo general y tres tipos de restricciones:

H₀: Los parámetros son los mismos para todos los individuos.

$$\hat{Y}_{rt}^0 = \alpha + X_{rt}\beta + \varepsilon_{rt} \quad \begin{matrix} r = 1, \dots, M \\ t = 1, \dots, T \end{matrix}$$

H₁: Las pendientes de la regresión son idénticas pero los términos independientes son diferentes.

$$\hat{Y}_{rt}^1 = \alpha_r + X_{rt}\beta + \varepsilon_{rt} \quad \begin{matrix} r = 1, \dots, M \\ t = 1, \dots, T \end{matrix}$$

H₂: Los términos independientes coinciden pero las pendientes son diferentes.

$$\hat{Y}_n^2 = \alpha + X_n \beta_r + \varepsilon_n \quad \begin{matrix} r = 1, \dots, M \\ t = 1, \dots, T \end{matrix}$$

H₃: Parámetros diferentes para cada individuo.

$$\hat{Y}_n^3 = \alpha_r + X_n \beta_r + \varepsilon_n \quad \begin{matrix} r = 1, \dots, M \\ t = 1, \dots, T \end{matrix}$$

Debido a que es muy poco frecuente tener en cuenta la hipótesis H₂ no consideraremos esa restricción y trabajaremos solamente con las demás. Nos referiremos ahora al modelo H₃ como el modelo sin restricciones, a H₁ como el modelo corregido por las medias de los individuos y a H₀ como el modelo completo.

Para realizar las estimaciones necesitamos hacer algunas suposiciones sobre el comportamiento de las perturbaciones, y en particular sobre el comportamiento de su matriz de Covarianzas.

En un primer subapartado se introduce la estimación del modelo en el caso más sencillo (con homocedasticidad y sin autocorrelación) para ver como se estimaría cada uno de los modelos restringidos.

A continuación se analiza la estimación de estos modelos en el caso general, desde el punto de vista teórico, para ir concretando a los principales casos particulares, que han sido tratados en la literatura existente.

2.1.2 Estimación del modelo (con homocedasticidad y sin autocorrelación).

En este caso vamos a suponer lo siguiente

$$\text{Var}(\varepsilon_{rt}) = \sigma^2, \quad \begin{matrix} r = 1, \dots, M \\ t = 1, \dots, T \end{matrix}$$

$$E(\varepsilon_{rt}, \varepsilon_{sv}) = 0, \quad \begin{matrix} s, r = 1, \dots, M \\ v, t = 1, \dots, T \end{matrix} \quad \text{con } r \neq s \text{ y } t \neq v.$$

$$E(\varepsilon_{rt}, \varepsilon_{rv}) = 0, \quad \begin{matrix} r = 1, \dots, M \\ v, t = 1, \dots, T \end{matrix} \quad \text{con } t \neq v.$$

Por consiguiente se está suponiendo que la varianza en cada periodo de tiempo y para cada entidad es constante. Además no existe información del pasado que no esté recogida en el modelo, así como tampoco existe influencia de las entidades entre sí.

El modelo sin restricciones lo podemos expresar como:

$$Y = W \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \varepsilon; \quad \text{Var}(\varepsilon) = \sigma^2 \cdot I_{MT}$$

Siendo:

$$W = \begin{pmatrix} W_1 \\ W_2 \\ \dots \\ W_M \end{pmatrix} = \begin{pmatrix} 1_T X_1 \\ 1_T X_2 \\ \dots \\ 1_T X_M \end{pmatrix},$$

X_r la matriz $T \times M$ de observaciones del individuo r -ésimo.

$$1_T = (1, 1, \dots, 1, 1)' \quad \text{con } T \text{ veces } 1.$$

Y el vector $MT \times 1$ de las observaciones de la variable explicada.

El estimador de mínimos cuadrados que se obtiene para el modelo sin restricción es:

$$\hat{\beta}_r = \left[(X_r - 1_r \bar{x}_r') (X_r - 1_r \bar{x}_r') \right]^{-1} \left[(X_r - 1_r \bar{x}_r') (Y_r - 1_r \bar{y}_r) \right]$$

$$\hat{\alpha}_r = \bar{y}_r - \bar{x}_r' \hat{\beta}_r \quad r = 1, 2, \dots, M$$

estimándose la varianza residual mediante la expresión:

$$\hat{\sigma}_{3\varepsilon}^2 = \sum_{r=1}^M \left[\frac{(Y_r - 1_r \bar{y}_r) (Y_r - 1_r \bar{y}_r) - \hat{\beta}_r' (X_r - 1_r \bar{x}_r') (Y_r - 1_r \bar{y}_r)}{(T-p-1)M} \right]$$

\bar{x}_r es el vector columna de medias aritméticas de la variable X , calculadas para el individuo r -ésimo, Y_r es el vector de observaciones de la variable explicada para el individuo r -ésimo, y \bar{y}_r la media de la variable explicada para el individuo r -ésimo.

Si se trabaja con el modelo corregido por las medias individuales se tiene una matriz:

$$W = \begin{bmatrix} 1_T & 0_T' & 0_T & X_1 \\ 0_T & 1_T & 0_T & X_2 \\ 0_T & 0_T & 1_T & X_M \end{bmatrix},$$

dando lugar a los estimadores de los coeficientes:

$$\hat{\beta} = \left[\sum_{r=1}^M \left[(X_r - 1_r \bar{x}_r') (X_r - 1_r \bar{x}_r') \right] \right]^{-1} \left[\sum_{r=1}^M \left[(X_r - 1_r \bar{x}_r') (Y_r - 1_r \bar{y}_r) \right] \right]$$

$$\hat{\alpha}_r = \bar{y}_r - \bar{x}_r' \hat{\beta}_r \quad r = 1, 2, \dots, M$$

La varianza residual se obtiene mediante la expresión:

$$\hat{\sigma}_{1\varepsilon}^2 = \sum_{r=1}^M \left[\frac{(Y_r - 1_r \bar{y}_r) (Y_r - 1_r \bar{y}_r) - \hat{\beta}_r' (X_r - 1_r \bar{x}_r') (Y_r - 1_r \bar{y}_r)}{(T-1)M-p} \right]$$

En el caso del modelo completo se tendría una matriz:

$$W = \begin{bmatrix} 1_T & 0_T & 0_T & X_1 & 0_{T \times p} & 0_{T \times p} \\ 0_T & 1_T & 0_T & 0_{T \times p} & X_2 & 0_{T \times p} \\ 0_T & 0_T & 1_T & 0_{T \times p} & 0_{T \times p} & X_M \end{bmatrix}$$

los estimadores que se obtendrían para los coeficientes serían:

$$\hat{\beta} = \left[\sum_{r=1}^M \left[(X_r - 1_T \bar{x}_r)' (X_r - 1_T \bar{x}_r) \right] \right]^{-1} \left[\sum_{r=1}^M \left[(X_r - 1_T \bar{x}_r)' (Y_r - 1_T \bar{y}_r) \right] \right]$$

$$\hat{\alpha} = \bar{y} - \bar{x}' \hat{\beta}$$

y para la varianza residual:

$$\hat{\sigma}_{0E}^2 = \sum_{r=1}^M \left[\frac{(Y_r - 1_T \bar{y}_r) (Y_r - 1_T \bar{y}_r)' - \hat{\beta}' (X_r - 1_T \bar{x}_r) (Y_r - 1_T \bar{y}_r)}{TM - (p+1)} \right]$$

La comparación entre estos tres estimadores permitirá averiguar cual de los modelos se ajusta mejor a los datos.

Esta idea la generalizaremos posteriormente para contrastar el modelo especificado en el caso no paramétrico.

2.1.3 Estimación del modelo general.

En este caso podemos considerar que la matriz de covarianzas de las perturbaciones es una matriz

$$\Omega = \begin{bmatrix} E(\mathcal{E}_{11}^2) & E(\mathcal{E}_{11} \mathcal{E}_{12}) & E(\mathcal{E}_{11} \mathcal{E}_{MT}) \\ E(\mathcal{E}_{12} \mathcal{E}_{11}) & E(\mathcal{E}_{12}^2) & E(\mathcal{E}_{12} \mathcal{E}_{MT}) \\ E(\mathcal{E}_{MT} \mathcal{E}_{11}) & E(\mathcal{E}_{MT} \mathcal{E}_{12}) & E(\mathcal{E}_{MT}^2) \end{bmatrix}$$

y en consecuencia no se presupone nada sobre éstas, por lo que pueden dar información de su pasado, e incluso pueden afectarse unas entidades a otras o variar su varianza, tanto a lo largo del tiempo como entre entidades. Se planteará el modelo general porque facilita comprender mejor los casos particulares, pero es bien sabido (Maddala (1971)) que el modelo no es identificable en este caso, si no se imponen restricciones en la varianza.

Todos los modelos implicados en la estimación en datos de panel se pueden expresar de la siguiente forma:

$$Y = WB + \varepsilon$$

siendo W una matriz que incluye los datos de las variables exógenas y B una matriz conteniendo los parámetros del modelo, salvo los de la varianza. Estas matrices variarán según las hipótesis que consideremos.

Mediante el método de los Mínimos Cuadrados Generalizados (MCG), el estimador de la matriz de parámetros B que se obtiene es (Novales (1993)):

$$B = (W' \Omega^{-1} W)^{-1} (W' \Omega^{-1} Y)$$

Según la forma que tome la matriz Ω tendremos diferentes formas para los estimadores de B . Por ejemplo, el caso estudiado en el apartado anterior se obtiene como consecuencia de suponer $\Omega = \sigma^2 I_{MT}$.

Normalmente, a la hora de hacer estas estimaciones, nos vamos a encontrar con el inconveniente de que los valores de la matriz Ω son desconocidos. Para estimar los valores de Ω no son suficientes las observaciones que obtengamos, ya que cada nueva observación implicará un número mayor de valores de Ω para estimar.

Por esa causa, en la práctica se considera que la matriz Ω sigue un modelo determinado que nos permita reducir el número de parámetros a estimar, volviéndose así factible la estimación del modelo.

Un ejemplo de esto, que engloba una gran parte de los modelos analizados, sería:

Modelo heterocedástico en sentido transversal y autorregresivo en sentido temporal.

Este modelo se caracteriza de la manera siguiente:

$$\begin{aligned} E(\varepsilon_{it}^2) &= \sigma_i^2 & u_{it} &\in N(0, \sigma_{u_i}^2) \\ E(\varepsilon_{it} \varepsilon_{jt}) &= 0 \quad (i \neq j) & \text{siendo: } \varepsilon_{it} &\in N\left(0, \frac{\sigma_{u_i}^2}{1 - \rho_i^2}\right) \\ \varepsilon_{it} &= \rho_i \varepsilon_{it-1} + u_{it} & E(\varepsilon_{it-1}, u_{it}) &= 0 \quad \forall i, j \end{aligned}$$

A partir de estas especificaciones se deduce que:

$$\begin{aligned} E(\varepsilon_{it} \varepsilon_{is}) &= \rho_i^{t-s} \sigma_i^2 \quad (t > s) \\ E(\varepsilon_{it} \varepsilon_{jis}) &= 0 \quad (i \neq j) \end{aligned}$$

La matriz Ω de este modelo será:

$$\Omega = \begin{pmatrix} \sigma_1^2 P_1 & 0_{TxT} & 0_{TxT} \\ 0_{TxT} & \sigma_2^2 P_2 & 0_{TxT} \\ 0_{TxT} & 0_{TxT} & \sigma_M^2 P_M \end{pmatrix} \quad \text{con } P_i = \begin{pmatrix} 1 & \rho_i & \rho_i^2 & \rho_i^{T-1} \\ \rho_i & 1 & \rho_i & \rho_i^{T-2} \\ \rho_i^{T-1} & \rho_i^{T-2} & \rho_i^{T-3} & \rho_i^{T-4} \\ \rho_i^{T-3} & \rho_i^{T-4} & \rho_i^{T-5} & 1 \end{pmatrix}$$

Haciendo una estimación de los parámetros α , β por MCO podemos obtener unos residuos estimados $\hat{\varepsilon}_{it}$. A partir de ellos se pueden calcular valores estimados para $\hat{\rho}_i$ y \hat{u}_{it} , $i=1, 2, \dots, M$; $t=1, 2, \dots, T$; y a partir de estos valores estimar $\hat{\sigma}_i^2$, con lo que tendríamos $\hat{\Omega}_i$ y a partir de ahí obtener los estimadores MCG del modelo que nos interesa.

2.2 INTRODUCCIÓN A LA ESTIMACIÓN NO PARAMÉTRICA Y SEMIPARAMÉTRICA EN LOS MODELOS ECONÓMICOS.

Cuando queremos conocer el comportamiento de una variable aleatoria partimos de una serie de datos observados, mediante los cuales intentaremos estimar una curva que nos describa el comportamiento teórico de la variable.

Las curvas de mayor interés para conocer ese comportamiento son:

Función de densidad: $f(x)$

Función de distribución: $F(x) = \int_{-\infty}^x f(t)dt$

La función de distribución se puede estimar directamente una vez conocida la función de densidad, por lo que centraremos el estudio en la estimación de la función de densidad.

Los métodos que existen para realizar esas estimaciones se pueden clasificar en dos grandes grupos: *paramétricos* y *no paramétricos*.

Los métodos paramétricos suponen que los datos provienen de una distribución caracterizada por un número finito de parámetros que estimaremos a partir de los datos, mientras que los métodos no paramétricos hacen solamente suposiciones muy generales e intentan estimar la forma de la función y su estructura.

2.2.1 Métodos paramétricos.

Históricamente, los métodos más empleados fueron los paramétricos, debido a que los métodos no paramétricos necesitan cálculos mucho más laboriosos para su obtención.

En los métodos paramétricos se dan en general dos enfoques, según se estudien a partir de la información que suministran los datos sobre la distribución de la variable o se tenga en cuenta las distancias geométricas de los errores.

En el primer caso tenemos:

Método de los Momentos. Fue formalizado por Pearson a finales del siglo XIX, siendo el primer método empleado para obtener la estimación de un parámetro de una distribución. La idea de este método consiste en que, si pretendemos estimar Θ , un vector $k \times 1$ de parámetros, que es función de k momentos de la distribución, sustituiremos el valor teórico desconocido de cada momento por el correspondiente momento muestral, de forma que se obtiene un sistema de k ecuaciones con k incógnitas y a partir de sus soluciones calcular un vector de parámetros Θ que sea función de los valores de la muestra.

Método de Máxima Verosimilitud. Introducido por Fisher (1922), teniendo más aceptación que el método de los momentos. Este método parte del principio de que ocurre lo más probable, por lo tanto, si tenemos una muestra y suponemos que procede de una variable con una determinada distribución, tomaremos como sus parámetros aquellos que hagan máxima la probabilidad de obtener los valores observados.

En el segundo caso está:

Método de los Mínimos Cuadrados. Es el primer método que se empleó en la estimación de los parámetros de una línea de regresión. Los primeros trabajos aparecidos sobre este método se deben a Legendre (1805). Años más tarde Gauss (1823) y Markov (1900) publican nuevos trabajos desarrollando este método. La idea de este método es la de escoger los parámetros que hagan menor el error cometido al considerar la función de densidad estimada como la verdadera. Busca por lo tanto que la diferencia entre los valores observados y los valores teóricos que nos proporciona la función de densidad estimada sea el menor posible.

2.2.2 Métodos no paramétricos.

Al aumentar la potencia de cálculo de los ordenadores, comenzaron a cobrar interés los métodos no paramétricos ya que, bajo unos supuestos previos muy pequeños, permiten realizar una estimación más exacta de la función de interés.

Consideraremos a partir de ahora que trabajamos con una muestra $\{X_1, X_2, \dots, X_n\}$ obtenida de la observación de una variable X de dimensión p , que tiene una función de densidad desconocida $f(x)$ con soporte S .

2.2.2.1 Método histograma

Fue el primer intento de estimación de una función de densidad de forma desconocida (John Grant (1662)) y representa la fuente básica de inspiración de la mayoría de los métodos de estimación no paramétrica que surgen, fundamentalmente, a partir de la década de los 50.

Este estimador realiza una partición del soporte de la función de densidad y, para

cada punto $x \in S$, la densidad estimada tomaría como valor la frecuencia relativa de los valores observados que caen dentro del intervalo de x , dividida entre la longitud del intervalo.

Para $p=1$, formalizaríamos esto de la siguiente manera:

Sea $P_n = \{A_j: j \in \mathbb{N}\}$ una partición de \mathbb{R} con $\|P_n\| = h$, y sea I_j la función indicadora de A_j , o sea:

$$I_j(x) = \begin{cases} 1 & \text{si } x \in A_j \\ 0 & \text{si } x \notin A_j \end{cases}$$

El estimador de la densidad sería:

$$\hat{f}_n(x) = (nh)^{-1} \sum_{i=1}^n \sum_{j=1}^{\infty} I_j(x) \cdot I_j(X_i)$$

2.2.2.2 Método kernel.

Rosemblat (1956) generaliza el método histograma de la manera siguiente:

$$\hat{f}_n(x) = \frac{\#\{X_i \in (x - h_n, x + h_n)\}}{2nh_n} = \frac{\hat{F}_n(x + h_n) - \hat{F}_n(x - h_n)}{2nh_n}$$

donde $\{X_1, X_2, \dots, X_n\}$ es la muestra de partida, \hat{F}_n la función de distribución empírica, $x \in S \subset \mathbb{R}$, $h_n > 0 \forall n$.

Este estimador también se puede expresar como:

$$\hat{f}_n(x) = \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right)$$

$$\text{siendo } K(u) = \begin{cases} \frac{1}{2} & \text{si } |u| \leq 1 \\ 0 & \text{si } |u| > 1 \end{cases}$$

Basándose en esa expresión Parzen (1962) considera estimadores del tipo

$$\hat{f}_n(x) = \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right)$$

donde a la función $K: \mathbb{R} \rightarrow \mathbb{R}$ se le suele llamar función Kernel (o núcleo).

Cacoullos (1965) extiende esta definición a estimadores p-dimensionales:

$$\hat{f}_n(x) = \frac{1}{nh_n^p} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n^p}\right)$$

siendo $K: \mathbb{R}^p \rightarrow \mathbb{R}$ un Kernel p-dimensional.

2.2.2.3 Método de los spline.

Sea $S=[a,b]$, dada una partición Δ en $[a,b]$, siendo $\{y_0, y_1, y_2, \dots, y_k\}$ los nudos de la partición. Una función *spline* respecto de esa partición es una función definida a trozos según los distintos intervalos de la partición. Si además verifica que es $n-1$ veces diferenciable en $[a,b]$ y que su restricción a $[y_i, y_{i+1}]$, $i=1,2,\dots,k-1$, es un polinomio de grado n , entonces se dice que la función es un *polinomio spline de grado n* respecto de la partición anterior.

En la estimación no paramétrica interesan especialmente los spline cúbicos de interpolación naturales. Estos son polinomios spline de interpolación de grado 3 que en los nudos de la partición toman los mismos valores que una función h a la que interpolan, y además las segundas derivadas en los extremos del intervalo $[a, b]$ son cero.

El interés de estos splines es debido a que se quiere que las estimaciones que realizamos no sean excesivamente bruscas, por lo que se les impone unas condiciones de suavidad.

La importancia de los splines cúbicos de interpolación naturales radica en que son la solución más parsimoniosa (con menor número de parámetros) a la condición de suavidad siguiente:

$$\min_s \int_a^b (s''(x))^2 dx$$

sujeito a que s sea dos veces diferenciable continuamente y $s(y_i) = h(y_i)$, con $i=0, 1, \dots, k$.

La primera aplicación de las funciones spline a la estimación no paramétrica fue hecha en Boneva, Kendall, Stefanov (1971).

Este método intenta ser una suavización del estimador histograma en el que se corrige su discontinuidad.

Boneva, Kendall, Stefanov (1971) sugieren como suavización del estimador

$$\hat{f}_n(x) = n^{-1} \sum_{i=1}^n \sum_{j=-\infty}^{\infty} I_{\{I_j\}}(x) \cdot I_{\{I_j\}}(X_i)$$

cuando $I_j = [j, j+1)$, con $j \in \mathbb{Z}$, una función spline h que sea solución del siguiente problema de minimización:

$$\min_h \int_{-\infty}^{\infty} h'(x)^2 dx$$

sujeto a: h diferenciable continuamente y

$$\int_j^{j+1} \hat{f}_n(x) dx = n^{-1} \sum_{i=1}^n I_{\{I_j\}}(X_i) = \int_j^{j+1} h(x) dx, \quad \forall x \in \mathbb{Z}$$

Berlinet (1979) da la siguiente expresión para los estimadores spline:

Si $\Delta = \{y_0, y_1, y_2, \dots, y_N\}$ es una partición en $[0, 1]$

$$\hat{f}_n(x) = n^{-1} \sum_{i=1}^n K_n(x, X_i)$$

siendo $K_n(x, u) = \frac{d}{dx} S_{\gamma_{u, \Delta}}(x)$

$$\text{donde } \gamma_u(x) = \begin{cases} 1 & \text{si } x \geq u \\ 0 & \text{si } x < u \end{cases}$$

$S_{\gamma_{u, \Delta}}(x)$ un spline cúbico de interpolación de γ_u ,

verificando alguna condición de suavidad.

2.2.2.4 Método de las interpolaciones.

Es tratado en los trabajos de Wahba (1971) y Wahba (1975). Sea $S=[0,1]$ el soporte de la función de densidad que queremos estimar y \hat{F}_n la función de distribución empírica de la muestra.

Sea $\Delta = \{y_0, y_1, y_2, \dots, y_k\}$ una partición de $[0,1]$, siendo $y_j = \frac{j}{k}$ $j=0, 1, \dots, k$.

Definimos el estimador de la densidad como:

$$\hat{f}_n(x) = \left(\hat{G}_n(x) \right)'(x)$$

siendo $\hat{G}_n(x)$ el polinomio de interpolación de \hat{F}_n en los nudos y_j , $j=0, 1, \dots, k$.

2.2.2.5 Método de máxima verosimilitud penalizado.

Este método intenta buscar dentro de un conjunto H de posibles funciones aquella que maximiza la probabilidad de obtener la muestra de partida, (en este caso maximiza la función de verosimilitud). Se introduce además un termino de penalización, con el fin de suavizar la estimación obtenida.

Montricher, Tapia, Thompson (1975) sugieren el intento de solucionar el siguiente problema de optimización:

$$\max_i \prod_{i=1}^n h(x_i) \cdot e^{-\phi(h)}$$

sujeto a: $h \in H$, $\int_a^b h(t)dt = 1$, y $h(t) \geq 0 \quad \forall t \in (a,b)$ en el que ϕ es un funcional $\phi: H \rightarrow \mathbb{R}$,

y H un espacio de Hilbert.

2.2.2.6 Método de los k-puntos próximos.

Este método fija los k puntos más próximos a x y estima $f(x)$ como la proporción que esos datos representan respecto del volumen que encierran al considerar la bola de centro x y radio la distancia de x al k -ésimo punto más próximo de la muestra.

Si $\{X_{(1)}^x, X_{(2)}^x, \dots, X_{(k)}^x\}$ representa la ordenación de la muestra en función de su distancia a x ; $V_{k_n}(x)$ es el volumen de la bola de centro x y radio la distancia de x a

$X_{(k_n)}^x$, con $k_n \rightarrow \infty$, $\frac{k_n}{n} \rightarrow 0$ cuando $n \rightarrow \infty$; Loftsgarden, Quesenberry (1965) proponen

como estimador:

$$\hat{f}_n(x) = \frac{k_n}{nV_{k_n}(x)}$$

que se puede escribir como:

$$\hat{f}_n(x) = \frac{\sum_{i=1}^n K_n \left(\frac{x - X_i}{\|x - X_{k_n}^x\|} \right)}{\|x - X_{k_n}^x\|^p}$$

$$\text{con } K(u) = \begin{cases} 1 & \text{si } \|u\| \leq 1 \\ 0 & \text{si } \|u\| > 1 \end{cases}$$

Moore, Jackel (1976) proponen como generalización estimadores similares pero con otro tipo de Kernels p -dimensionales.

Devroye (1976) propone la siguiente generalización:

Si $\{K_n^1\}$ y $\{K_n^2\}$ son dos sucesiones verificando:

i) $K_n^1 \rightarrow \infty$ cuando $n \rightarrow \infty$

ii) $1 \leq K_n^1 \leq K_n^2 \leq n$, con $\frac{K_n^2}{n} \rightarrow 0$

y $\gamma_n = \{\gamma_{K_n^1}, \dots, \gamma_{K_n^2}\}$ una distribución de probabilidad.

Define como estimador:

$$\hat{f}_n(x) = \sum_{j=K_n}^{K_n^2} \frac{j}{nV_j(x)} \gamma_j$$

2.2.2.7 Método de las series ortogonales.

La introducción de esta técnica en la estimación no paramétrica es debida a Cencov (1962).

En ese trabajo considera S el soporte de f ; ω una función de peso medible sobre S y no negativa; y el espacio de Hilbert $L^2(S)$, con el producto interior:

$$\langle g_1, g_2 \rangle = \int_S g_1(x) \cdot g_2(x) \cdot \omega(x) dx,$$

y una base ortonormal $\{\psi_j\}_{j=1}^{\infty}$.

Si consideramos $f \in L^2(S)$, entonces la podemos expresar como:

$$\sum_{k=1}^{\infty} a_k(x) \psi_k(x)$$

siendo $a_k(x) = \int_S \psi_k(x) f(x) \cdot \omega(x) dx = E[\psi_k \cdot \omega]$ el coeficiente de Fourier.

Teniendo esto en cuenta, si queremos estimar la función f , necesitaremos estimar sus coeficientes de Fourier, o sea

$$\hat{a}_k = \frac{1}{n} \sum_{j=1}^n \psi_k(X_j) \omega(X_j) \quad k=1, 2, \dots, \infty$$

con lo cual se tendría:

$$\hat{f}_n(x) = \sum_{k=1}^{\infty} \hat{a}_k(x) \psi_k(x)$$

Ante la imposibilidad de estimar los infinitos coeficientes de Fourier distintos de cero, es necesario limitarse a calcular solamente un nº determinado de ellos, $N(n)$, que estará en función del tamaño de la muestra, n . Este valor ejerce la misma función de suavización que el parámetro ventana en la estimación Kernel.

El estimador sería:

$$\hat{f}_n(x) = \frac{1}{n} \sum_{j=1}^n \sum_{k=1}^{N(n)} \psi_k(x) \psi_k(X_j) \cdot \omega(X_j)$$

Otra posible variación está en la función de peso ω a emplear.

2.2.2.8 Estimadores delta.

Los estimadores delta son un intento de construir un formato general para varios estimadores no paramétricos.

Estos estimadores aparecen tratados en González (1982) y Härdle (1990)

Así tendríamos que la formulación general para la función de densidad sería:

$$\hat{f}_n(x) = n^{-1} \sum_{i=1}^n \delta_m(x, X_i)$$

con $\{\delta_m(x, X_i)_{m=m(n) \rightarrow \infty}\}$ una sucesión de funciones que indican el peso del dato X_i , y que se definen como:

$$\delta_m: \mathbb{R}^p \times \mathbb{R}^p \rightarrow \mathbb{R}$$

verificando: $\delta_m(x, u) \geq 0 \quad \forall x, u \in \mathbb{R}^p$

$$\int \delta_m(x, u) du = 1, \quad \forall u \in \mathbb{R}^p$$

A esta formulación se adaptan algunos de los métodos antes indicados:

Para el **Histograma**, $\delta_m(x, u) = h^{-1} \sum_{j=1}^{\infty} I_j(x) \cdot I_j(u)$

Para el método **Kernel**, $\delta_m(x, u) = h^{-1} K\left(\frac{x-u}{h}\right)$.

Para las **Series Ortogonales**, $\delta_m(x, u) = \sum_{k=1}^{N(n)} \psi_k(x) \psi_k(u) \cdot \omega(u)$

Para un estudio con más detalle de los métodos de estimación no paramétrica pueden consultarse, por ejemplo, González (1982), Silverman (1986), Härdle (1990), Prakasa (1983).

2.2.3 Estimación paramétrica mediante suavización no paramétrica previa.

La base esencial de la teoría de la regresión paramétrica consiste en la estimación de un parámetro Θ p-dimensional en el modelo

$$Y = A^t(X)\Theta + \varepsilon$$

en el que (X, Y) es un vector aleatorio $q+1$ -dimensional, ε una variable aleatoria de media cero y $A: \mathbb{R}^q \rightarrow \mathbb{R}^p$ un funcional arbitrario.

Disponemos de una muestra de partida $\{(X_1, Y_1), \dots, (X_n, Y_n)\}$ de la que obtendremos la información para estimar Θ .

Este modelo es general, pudiéndose adaptar a él los distintos modelos de regresión empleados en la práctica. Por ejemplo:

- $A^t(x) = (1, x, \dots, x^{p-1})$, con $x \in \mathbb{R}$, sería el modelo de regresión polinómica;
- $A^t(X) = (x_1, x_2, \dots, x_p)$ es el modelo de regresión múltiple;
- si construimos el vector (X, Y) con $X = (z_1, z_2, \dots, z_p)$ e $Y = z_{p+1}$, donde $\{z_i\}_{i \in \mathbb{Z}}$ es un modelo estacionario univariante, se obtiene un modelo de autorregresión de orden p .
- Finalmente, según X sea aleatorio o determinístico tendremos los modelos de correlación o diseño aleatorio, y los modelos de diseño fijo o experimental, respectivamente.

El método más desarrollado en este contexto a lo largo de este siglo fue el de los mínimos cuadrados, que intenta minimizar la norma euclídea de los errores aleatorios, o sea intenta minimizar el funcional:

$$\psi_0(\Theta) = \sum_{i=1}^n (Y_i - A'(X_i)\Theta)^2$$

Cuando los errores $\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}$ siguen una distribución normal el estimador de mínimos cuadrados es el estimador de máxima verosimilitud y sus propiedades son conocidas de manera exacta.

Sin embargo, en la práctica, los tests de normalidad de los residuos nos indican con frecuencia la no normalidad de los errores. En este tipo de situaciones los estimadores de mínimos cuadrados pierden propiedades de eficiencia y consistencia. Por ejemplo, en el trabajo de Ruppert, Carroll (1980) se comprueba mediante simulación resultados de gran ineficiencia para los antedichos estimadores en situaciones en las que el error sigue una distribución con colas más pesadas que la Normal. Otros trabajos en los que se ven los problemas de la falta de normalidad de los residuos son los de Drygas (1976) y Lai, Robbins, Wei (1978).

Consideraciones de ese tipo llevaron a la busca de estimadores alternativos. Entre estas posibilidades estaría la de emplear un estimador no paramétrico como fuente de información inicial, es decir, si disponemos de un estimador no paramétrico piloto $\hat{\alpha}_n(x)$ y del modelo paramétrico $Y=E[Y/X=x]=A'(X)\Theta$, es posible emplear $\hat{\alpha}_n(x)$ definiendo como estimador de Θ al valor que minimiza algún tipo de distancia funcional entre $\hat{\alpha}_n(x)$ y la función de regresión paramétrica del modelo.

Dentro de las posibles distancias, la única abordada hasta el momento es la definida como

$$\psi_1(\Theta) = \int (\hat{\alpha}_n(x) - A'(X_i)\Theta)^2 d\Omega_n(x)$$

donde $\Omega_n(x)$ es una función de ponderación construida a partir de la muestra.

En el caso particular de

$$\hat{\alpha}_n(x) = \sum_{i=1}^n Y_i I_{\{X_i=x\}} \quad \Omega_n(x) = \frac{\#\{X_i \leq x\}}{n}$$

($\Omega_n(x)$ es la función de distribución empírica de la muestra)

este estimador se convierte en el de mínimos cuadrados.

En los trabajos realizados hasta ahora con esta técnica los principales estimadores piloto empleados fueron los estimadores delta:

$$\hat{\alpha}_n(x) = \frac{\sum_{i=1}^n Y_i \delta_m(x, X_i)}{\sum_{r=1}^n \delta_m(x, X_r)}$$

donde $\{\delta_m: \mathbb{R}^q \times \mathbb{R}^q \rightarrow \mathbb{R}\}_{m=m(n) \rightarrow \infty}$ es una sucesión de funciones medibles, mientras que la función de ponderación es:

$$\Omega_n(x) = \int_{-\infty}^{\infty} \hat{f}_n(t) dt = n^{-1} \sum_{i=1}^n \int_{-\infty}^{\infty} \delta_m(t, X_i) dt$$

2.2.3.1 La recta de regresión.

Cuando la muestra de partida $\{(X_1, Y_1), \dots, (X_n, Y_n)\}$ es bidimensional y $A^1(x) = (1, x)$ estamos en el caso más simple de la regresión, la recta de regresión. Este modelo es del tipo:

$$Y = \theta_1 + \theta_2 X + \varepsilon$$

del que supondremos que ε es una variable aleatoria de media cero y varianza σ^2 .

Los primeros trabajos en los que se aborda un estudio inicial de los nuevos estimadores aplicados a este modelo son Faraldo (1982), donde realiza algunas simulaciones ilustrativas de los mismos, y Faraldo (1984), donde realiza un primer estudio teórico probando resultados relativos a su consistencia y normalidad asintótica tomando como estimadores piloto no paramétricos los estimadores delta.

En este caso los estimadores $\hat{\theta}_1$ y $\hat{\theta}_2$ resultan:

$$\hat{\theta}_1 = \frac{\bar{Z} \cdot \bar{V} - \bar{T} \cdot \bar{W}}{\bar{V} \cdot \bar{D} - \bar{W}^2} \quad \text{y} \quad \hat{\theta}_2 = \frac{\bar{T} \cdot \bar{D} - \bar{Z} \cdot \bar{W}}{\bar{V} \cdot \bar{D} - \bar{W}^2}$$

donde \bar{Z} , \bar{V} , \bar{W} , \bar{T} , \bar{D} son las medias de las variables:

$$Z_i = \int_S Y_i \delta_m(x, X_i) dx, \quad V_i = \int_S x^2 \delta_m(x, X_i) dx, \quad W_i = \int_S x \delta_m(x, X_i) dx$$

$$T_i = \int_S Y_i x \delta_m(x, X_i) dx, \quad \text{y} \quad D_i = \int_S \delta_m(x, X_i) dx, \quad i = 1, 2, \dots, n$$

siendo S un conjunto abierto conteniendo al soporte de f (la densidad de X), que se supone existente.

Faraldo, González (1987) prolongan ese estudio teórico probando propiedades asintóticas bajo las siguientes condiciones de los estimadores:

$$i) \quad \hat{\theta}_1 \xrightarrow{P} \theta_1, \quad \hat{\theta}_2 \xrightarrow{P} \theta_2, \quad \text{siempre que } S \text{ sea compacto, } E(Y^2) < \infty \text{ y}$$

$$m = m(n) \rightarrow \infty$$

$$ii) \quad \hat{\theta}_1 \xrightarrow{c.s.} \theta_1, \quad \hat{\theta}_2 \xrightarrow{c.s.} \theta_2, \quad \text{siempre que } S \text{ sea compacto, } Y \text{ acotada con}$$

probabilidad 1 y $m=m(n) \rightarrow \infty$

$$\text{iii) } \sqrt{n} \left[\begin{pmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \end{pmatrix}' - (\theta_1, \theta_2)' \right] \xrightarrow{d} N \left[(0,0)', B \right] \quad \text{con } B = \sigma^2 [E\{(1,x)'(1,x)\}]$$

siempre que S sea compacto y $E(|Y^{2+\gamma}|) < \infty$ para algún $\gamma > 0$ y $\sqrt{n} \cdot m \rightarrow 0$.

Sobre la sucesión $\{\delta_m\}$ se exigen las siguientes condiciones axiomáticas de tipo general:

$$\text{a) } \sup_u \delta_m(x,u) = O(m) \quad \forall x \in \mathbb{R}$$

$$\text{b) } \delta_m(x,t) = \delta_m(t,x) \text{ y } \delta_m(x,t) \geq 0 \quad \forall (x,t) \in \mathbb{R}^2$$

$$\text{c) } \int \delta_m(x,u) du = 1 \quad \forall x \in \mathbb{R} \text{ con } \delta_m(x,t) = 0 \text{ si } |x-t| > c\varepsilon_n \quad (\varepsilon_n = O(\frac{1}{m})) \text{ y } \varepsilon_n \rightarrow 0 \text{ si } n \rightarrow \infty$$

El **Histograma** o el **Kernel** con soporte compacto son ejemplos de estimadores en los que sus funciones δ_m verifican las condiciones anteriores.

Estos estimadores $\hat{\theta}$ pueden heredar las malas propiedades que presentan los estimadores piloto no paramétricos iniciales en situaciones del tipo $S=[a,b]$, ya que los estimadores no paramétricos presentan ineficiencias en las colas $[a, a+\varepsilon_n]$ y $[b-\varepsilon_n, b]$, y estas ineficiencias son trasladadas a estos estimadores $\hat{\theta}$. Sin embargo,

en una gran generalidad de las situaciones se producen estimaciones más eficientes que la mínima cuadrática (Faraldo, González (1987)).

Un estimador piloto no paramétrico que elimina el efecto frontera en los estimadores

$\hat{\theta}$ es el estudiado por Stute (1984) y se corresponde a un estimador del tipo de los **k-puntos próximos**:

$$\hat{\alpha}_n(x) = \frac{1}{n\mathcal{E}_n} \sum_{i=1}^n Y_i K \left[\frac{F_n(x) - F_n(X_i)}{\mathcal{E}_n} \right]$$

teniendo como función de ponderación la función de distribución empírica de la muestra $\hat{\Omega}_n = F_n(x)$. Estos estimadores son estudiados en Stute, González (1990) y González (1990).

2.2.3.2 Modelo de regresión múltiple: diseño fijo.

Sea el modelo:

$$Y = A'(X)\Theta + \varepsilon$$

en el que ε es una variable aleatoria de media cero y varianza σ^2 . A una función definida en un compacto C p -dimensional y Θ el parámetro a estimar.

Supongamos que disponemos de una muestra $\{(x_i, Y_i)\}$ $i=1, 2, \dots, n$, donde $\{x_1, \dots, x_n\}$ con $x_i \in C$, $i=1, 2, \dots, n$, representa el diseño prefijado de antemano. Debido a que los datos $\{x_i\}$ suelen ser controlados por el investigador, tienden a estar equiespaciados en C , por eso tomando como función de ponderación la distribución uniforme en C se definirían como nuevos estimadores de Θ los que optimizan el funcional :

$$\psi_1(\Theta) = \int (\hat{\alpha}_n(x) - A'(X_i)\Theta)^2 dx$$

donde $\hat{\alpha}_n(x)$ es un estimador no paramétrico de la función de regresión.

Un formato general para $\hat{\alpha}_n$ fue introducido por Georgiev (1983, 1985):

$$\hat{\alpha}_n(x) = \sum_{i=1}^n \omega_m(x) Y_i, x \in C \subset \mathbb{R}$$

conteniendo numerosos casos particulares ya tratados en la literatura, como los de Priestley, Chao (1972), Benedetti (1977), Gasser, Muller (1979).

Bajo la suposición general de existencia de inversa, el estimador general $\hat{\Theta}$ viene dado por:

$$\hat{\Theta}_n = \left[\int_C A(X) A'(X) dx \right]^{-1} \left[\int_C \hat{\alpha}_n(x) A(X) dx \right]$$

En González (1988) aparecen resultados de consistencia y normalidad asintótica para este estimador, bajo condiciones muy poco restrictivas.

2.2.3.3 Modelo de regresión múltiple: diseño aleatorio.

En este modelo partimos de una muestra $\{(X_1, Y_1), \dots, (X_n, Y_n)\}$ del vector $(p+1)$ -dimensional (X, Y) , y el funcional que nos da la forma de la regresión será:

$$A(X) = A((x_1, \dots, x_p)^t) = (x_1, \dots, x_p)^t$$

con lo cual el modelo de regresión será:

$$Y = A^t(x) \Theta = x_1 \theta_1 + \dots + x_p \theta_p$$

Cristobal, Faraldo, González (1987) proponen como estimadores piloto no paramétricos los estimadores delta ya vistos antes. Entonces, suponiendo que la función de distribución es absolutamente continua, y exigiendo para las funciones

$\{\delta_m\}$ las siguientes condiciones:

- a) $\sup_u \delta_m(x,u) = O(m^p) \forall x \in \mathbb{R}^p$
- b) $\delta_m(x,t) = \delta_m(t,x)$ y $\delta_m(x,t) \geq 0 \forall (x,t) \in \mathbb{R}^{2p}$
- c) $\int \delta_m(x,u) du = 1 \forall x \in \mathbb{R}^p$ con $\delta_m(x,t) = 0$ si $\|x-t\| > c\varepsilon_n$ ($\varepsilon_n = O(\frac{1}{m})$) y $\varepsilon_n \rightarrow 0$ si $n \rightarrow \infty$

demuestran lo siguiente:

i) Si $E(Y^4) < \infty$ entonces $\hat{\Theta} \xrightarrow{c.s.} \Theta$

ii) Si $E(\|(X,Y)\|^2) < \infty$ y si tiene $E(|Y|^{2+\gamma}) < \infty$ con $\gamma > 0$, entonces:

$$\sqrt{n}(\hat{\Theta} - \Theta) \xrightarrow{d} N(0, \sigma^2(E[X'X])^{-1})$$

Igual que en el caso de la recta de regresión, estos estimadores son asintóticamente eficientes, teniendo un error cuadrático medio inferior al de mínimos cuadrados.

En el trabajo de Cristobal, Faraldo, González (1987) también se demuestra que los estimadores Ridge (o riscal), introducidos por Hoerl, Kennard (1970) para la corrección del E.C.M en las variables con problemas de multicolinealidad, son un caso particular del método que estamos estudiando.

Usando estimadores piloto no paramétricos recursivos, tanto para la regresión como para la función de ponderación, podemos construir con este método unos estimadores recursivos, como se puede ver en González (1988).

2.2.3.5 Modelo de datos censurados.

Los modelos estadísticos con muestras censuradas son de especial interés en ciencias experimentales. Se corresponden, en general, con el estudio de T , la duración aleatoria de un proceso, con la particularidad de que factores ajenos a ese proceso y aleatorios, que representaremos con una variable C , pueden obligar a interrumpir el proceso antes de su final.

En Cadarso (1990) se considera el modelo para datos censurados:

$$T = A'(z_i)\Theta + \varepsilon_i$$

con $E(\varepsilon_i) = 0$ y $z_i \in C$, C compacto, $(i=1, \dots, n)$; suponiendo que existe una censura aleatoria por la derecha para la variable aleatoria T , condicionalmente independiente respecto de la covariable z .

Ahora la muestra de partida es:

$$\{(Y_1, \delta_1, z_1), \dots, (Y_n, \delta_n, z_n)\}$$

con $Y_i = \min\{T_i, C_i\}$, $\delta_i = I_{\{T_i \leq C_i\}}$ y T_i condicionalmente independiente a C_i ($i=1, 2, \dots, n$).

De manera análoga a los modelos anteriores, el estimador se obtendría de minimizar la expresión:

$$\psi_1(\Theta) = \int (\hat{\alpha}_n(z) - A'(z)\Theta)^2 d\Omega_n(z)$$

En base a la muestra inicial con datos censurados y tomando como Ω_n la distribución empírica de la muestra sobre las z_i , el estimador obtenido es el siguiente:

$$\hat{\Theta} = \left[\sum_{j=1}^n A(z_j) A'(z_j) \right]^{-1} \left[\sum_{i=1}^n A(z_i) \hat{\alpha}_n(x) \right]$$

comprobando además su normalidad asintótica.

2.2.3.6 Modelo de datos de panel.

Los modelos de datos de panel se corresponden, como ya hemos visto antes, con la observación de una variable en una serie de individuos o entidades a través de varios periodos de tiempo.

El modelo correspondiente parte de una muestra

$$\{(X_{rt}, Y_{rt})\} \quad \begin{matrix} r = 1, 2, \dots, M \\ t = 1, 2, \dots, T \end{matrix}$$

en la que (X, Y) es una variable aleatoria $(p+1)$ -dimensional, M el número de individuos o entidades observadas, y T el número de periodos de tiempo que dura la observación; X es la variable exógena p -dimensional e Y la variable explicada.

El modelo sobre el que se realiza este trabajo se considera con q parámetros, $0 \leq q \leq p$, que pueden variar a través de los individuos mientras que los otros $p-q$ se consideran constantes.

La forma que tomará el modelo será:

$$Y_{rt} = A'(X_{rt})\Theta + \varepsilon_{rt}$$

Donde:

- $\Theta = (\theta_1^1, \theta_2^1, \dots, \theta_q^1, \theta_1^2, \dots, \theta_1^M, \theta_{q+1}, \dots, \theta_p)'$ será un vector de parámetros de

$$\text{dimensión } c = (M-1)q + p = Mq + (p-q),$$

- $A: \mathbb{R}^c \rightarrow \mathbb{R}^p$ será el funcional:

$$A[(x^1, x^2, \dots, x^p)^t] = (I^1(x)x^1, I^1(x)x^2, \dots, I^1(x)x^q, I^2(x)x^1, \dots, I^M(x)x^q, x^{q+1}, \dots, x^p)^t$$

Con I^r , $r=1,2,\dots,M$ la función indicadora de la muestra r -ésima

- ε_{rt} una variable aleatoria verificando:

$$E(\varepsilon_{rt})=0, \quad V(\varepsilon_{rt})=\sigma^2, \quad \text{Cov}(\varepsilon_{rt}, \varepsilon_{su}) = \begin{cases} \sigma_{r,ru} & r = s, t \neq u \\ 0 & r \neq s \end{cases}$$

El estimador del vector de parámetros Θ , saldrá de la minimización del funcional:

$$\psi_1(\Theta) = \int (\hat{\alpha}_n(x) - A'(X_i)\Theta)^2 d\Omega_n(x)$$

por lo que la forma de ese estimador vendrá dada por la expresión:

$$\hat{\Theta} = \left[\int A(X)A'(X)d\Omega_n(x) \right]^{-1} \cdot \left[\int \hat{\alpha}_n(x)A(X)d\Omega_n(x) \right]$$

Que se obtiene fácilmente derivando el funcional ψ_1 e igualando a cero.

Empleando como estimadores no paramétricos previos los estimadores delta, y bajo condiciones muy generales, se demuestran en este trabajo la consistencia y la normalidad asintótica del estimador estudiado para el modelo de datos de panel.

2.3 ESTIMACIÓN SEMIPARAMÉTRICA CON DATOS DE PANEL.

2.3.1 Definición del estimador.

Se pretende estimar un modelo de regresión múltiple

$$Y = X'\Theta + \varepsilon$$

donde Y es la variable explicada,
 X un vector p -dimensional de variables exógenas,
 $\Theta = (\theta_1, \theta_2, \dots, \theta_p)'$ un vector p -dimensional de parámetros, y
 ε una variable aleatoria de media cero.

La particularidad que va a tener es que será estimado partiendo de una muestra

$$\{(X_{rt}, Y_{rt})\} \quad \begin{matrix} r = 1, 2, \dots, M \\ t = 1, 2, \dots, T \end{matrix}$$

procedente de la observación de M individuos durante T periodos de tiempo. Por este motivo, puede resultar que q parámetros, $0 \leq q \leq p$, varíen a través de los individuos, mientras los $p-q$ restantes permanecen constantes.

Teniendo todo esto en cuenta, el modelo tomará la forma siguiente:

$$Y_{rt} = A^t(X_{rt})\Theta + \varepsilon_{rt}$$

Donde:

- $\Theta = (\theta_1^1, \theta_2^1, \dots, \theta_q^1, \theta_1^2, \dots, \theta_q^2, \dots, \theta_q^M, \theta_{q+1}, \dots, \theta_p)$ será un vector de parámetros de dimensión $c = (M-1)q + p = Mq + (p-q)$,
- $A: \mathbb{R}^c \rightarrow \mathbb{R}^p$ será el funcional:

$$A[(x^1, x^2, \dots, x^p)'] = (I^1(x)x^1, I^1(x)x^2, \dots, I^1(x)x^q, I^2(x)x^1, \dots, I^M(x)x^q, x^{q+1}, \dots, x^p)'$$

Con I^r , $r=1,2,\dots,M$ la función indicadora de la muestra r -ésima

- ε_{rt} una variable aleatoria verificando:

$$E(\varepsilon_{rt})=0, \quad V(\varepsilon_{rt})=\sigma^2, \quad Cov(\varepsilon_{rt}, \varepsilon_{su}) = \begin{cases} \sigma_{r, su} & r = s, t \neq u \\ 0 & r \neq s \end{cases}$$

El estimador que emplearemos va a ser una extensión a datos de panel del propuesto en Faraldo, González (1987) para la regresión lineal y que se introducía en el *Apartado 2.2.3.1.*

Este estimador será el que minimice el siguiente funcional:

$$\psi_1(\Theta) = \int (\hat{\alpha}_n(x) - A'(X)\Theta)^2 d\Omega_n(x)$$

siendo $\hat{\alpha}_n(x)$ un estimador no paramétrico de $\alpha(x)$, $\Omega_n(x)$ una función de ponderación construida a partir de la muestra y $n=M \cdot T$ el número de datos de la muestra.

En este trabajo el estimador no paramétrico que usaremos será:

$$\hat{\alpha}_n(x) = \frac{\sum_{m=1}^M \sum_{t=1}^T Y_{tm} \delta_s(x, X_{tm})}{\sum_{m=1}^M \sum_{t=1}^T \delta_s(x, X_{tm})}$$

mientras que la función de ponderación será:

$$\Omega_n(x) = \int_{-\infty}^x \hat{f}_n(u) du = \frac{\sum_{m=1}^M \sum_{t=1}^T \int_{-\infty}^x \delta_s(u, X_{tm}) du}{M \cdot T}$$

En los que \hat{f}_n es un estimador no paramétrico de la densidad de X , y

$\{\delta_s: \mathbb{R}^p \times \mathbb{R}^p \rightarrow \mathbb{R}\}_{s=s(n) \rightarrow \infty}$ una sucesión de funciones medibles.

Entonces, derivando e igualando a cero el funcional ψ , obtendremos la siguiente expresión para el estimador:

$$\hat{\Theta} = \left[\int A(X) A'(X) d\Omega_n(x) \right]^{-1} \cdot \left[\int \hat{\alpha}_n(x) A(X) d\Omega_n(x) \right] = C^{-1} D$$

$$\text{siendo } C = \int A(X) A'(X) d\Omega_n(x); D = \int \hat{\alpha}_n(x) A(X) d\Omega_n(x)$$

Ahora, desarrollando esa expresión obtendremos:

$$\int \hat{\alpha}_n(x) A(X) d\Omega_n(x) = \int \hat{\alpha}_n(x) \begin{pmatrix} I^1(x) x^1 \\ I^1(x) x^2 \\ \vdots \\ I^1(x) x^q \\ I^2(x) x^1 \\ I^2(x) x^2 \\ \vdots \\ I^M(x) x^q \\ x^{q+1} \\ \vdots \\ x^p \end{pmatrix} d\Omega_n(x) = \begin{pmatrix} \int \hat{\alpha}_n(x) I^1(x) x^1 d\Omega_n(x) \\ \int \hat{\alpha}_n(x) I^1(x) x^2 d\Omega_n(x) \\ \vdots \\ \int \hat{\alpha}_n(x) I^1(x) x^q d\Omega_n(x) \\ \int \hat{\alpha}_n(x) I^2(x) x^1 d\Omega_n(x) \\ \int \hat{\alpha}_n(x) I^2(x) x^2 d\Omega_n(x) \\ \vdots \\ \int \hat{\alpha}_n(x) I^M(x) x^q d\Omega_n(x) \\ \int \hat{\alpha}_n(x) x^{q+1} d\Omega_n(x) \\ \vdots \\ \int \hat{\alpha}_n(x) x^p d\Omega_n(x) \end{pmatrix}$$

Como:

$$\hat{\alpha}_n(x) d\Omega_n(x) = \frac{\sum_{m=1}^M \sum_{t=1}^T Y_{tm} \delta_s(x, X_{tm}) dx}{T \cdot M} \Rightarrow$$

$$\Rightarrow \int \hat{\alpha}_n(x) A(X) d\Omega_n(x) = \begin{pmatrix} \frac{\sum_{m=1}^M \sum_{l=1}^T \int I^1(x) x^1 Y_{lm} \delta_s(x, X_{lm}) dx}{T \cdot M} \\ \frac{\sum_{m=1}^M \sum_{l=1}^T \int I^M(x) x^q Y_{lm} \delta_s(x, X_{lm}) dx}{T \cdot M} \\ \frac{\sum_{m=1}^M \sum_{l=1}^T \int x^{q+1} Y_{lm} \delta_s(x, X_{lm}) dx}{T \cdot M} \\ \frac{\sum_{m=1}^M \sum_{l=1}^T \int x^p Y_{lm} \delta_s(x, X_{lm}) dx}{T \cdot M} \end{pmatrix} = \begin{pmatrix} \frac{\sum_{m=1}^M \sum_{l=1}^T T_{1,lm}^1}{T \cdot M} \\ \frac{\sum_{m=1}^M \sum_{l=1}^T T_{q,lm}^M}{T \cdot M} \\ \frac{\sum_{m=1}^M \sum_{l=1}^T T_{q+1,lm}}{T \cdot M} \\ \frac{\sum_{m=1}^M \sum_{l=1}^T T_{p,lm}}{T \cdot M} \end{pmatrix} = \begin{pmatrix} \bar{T}_1 \\ \bar{T}_q^M \\ \bar{T}_{q+1} \\ \bar{T}_p \end{pmatrix}$$

Siendo $T_{i,lm}^r = \int x^i I^r(x) Y_{lm} \delta_s(x, X_{lm}) dx$; $T_{i,lm} = \int x^i Y_{lm} \delta_s(x, X_{lm}) dx$.

Operando de la misma manera tengo: $\int A(X) A'(X) d\Omega_n(x) =$

$$= \begin{pmatrix} \bar{V}_{11}^1 & \bar{V}_{1q}^1 & 0 & 0 & 0 & 0 & \bar{V}_{1q+1}^1 & \bar{V}_{1p}^1 \\ \bar{V}_{q1}^1 & \bar{V}_{qq}^1 & 0 & 0 & 0 & 0 & \bar{V}_{qq+1}^1 & \bar{V}_{qp}^1 \\ 0 & 0 & \bar{V}_{11}^2 & \bar{V}_{1q}^2 & 0 & 0 & \bar{V}_{1q+1}^2 & \bar{V}_{1p}^2 \\ 0 & 0 & \bar{V}_{q1}^2 & \bar{V}_{qq}^2 & 0 & 0 & \bar{V}_{qq+1}^2 & \bar{V}_{qp}^2 \\ 0 & 0 & 0 & 0 & \bar{V}_{11}^M & \bar{V}_{1q}^M & \bar{V}_{1q+1}^M & \bar{V}_{1p}^M \\ 0 & 0 & 0 & 0 & \bar{V}_{q1}^M & \bar{V}_{qq}^M & \bar{V}_{qq+1}^M & \bar{V}_{qp}^M \\ \bar{V}_{q+11}^1 & \bar{V}_{q+1q}^1 & \bar{V}_{q+11}^2 & \bar{V}_{q+1q}^2 & \bar{V}_{q+11}^M & \bar{V}_{q+1q}^M & \bar{V}_{q+1q+1}^M & \bar{V}_{q+1p}^M \\ \bar{V}_{p1}^1 & \bar{V}_{pq}^1 & \bar{V}_{p1}^2 & \bar{V}_{pq}^2 & \bar{V}_{p1}^M & \bar{V}_{pq}^M & \bar{V}_{pq+1}^M & \bar{V}_{pp}^M \end{pmatrix}$$

Siendo $V_{ij,lm}^r = \int x^i x^j I^r(x) Y_{lm} \delta_s(x, X_{lm}) dx$; $V_{ij,lm} = \int x^i x^j Y_{lm} \delta_s(x, X_{lm}) dx$; y

teniendo en cuenta que $I^r(x) \cdot I^s(x) = 0$ si $r \neq s$.

La relación entre los elementos \bar{V} y los c_{ij} es la siguiente:

$$\bar{V}_{jl}^i = c_{(i-1)q+j;(i-1)q+l} \quad \text{si } \begin{matrix} i = 1, 2, \dots, M \\ j = 1, 2, \dots, q \\ l = 1, 2, \dots, q \end{matrix}$$

$$\bar{V}_{jl}^i = c_{(i-1)q+j;(M-1)q+l} \quad \text{si } \begin{matrix} i = 1, 2, \dots, M \\ j = 1, 2, \dots, q \\ l = q+1, \dots, p \end{matrix}$$

$$\bar{V}_{jl}^i = c_{(M-1)q+j;(i-1)q+l} \quad \text{si } \begin{matrix} i = 1, 2, \dots, M \\ j = q+1, \dots, p \\ l = 1, 2, \dots, q \end{matrix}$$

$$\bar{V}_{jl}^i = c_{(M-1)q+j;(M-1)q+l} \quad \text{si } \begin{matrix} j = q+1, \dots, p \\ l = q+1, \dots, p \end{matrix}$$

La relación entre los elementos \bar{T} y los d_j es la siguiente:

$$\bar{T}_j^i = d_{(i-1)q+j} \quad \text{si } \begin{matrix} i = 1, 2, \dots, M \\ j = 1, 2, \dots, q \end{matrix}$$

$$\bar{T}_j^i = d_{(M-1)q+j} \quad \text{si } j = q+1, \dots, p$$

Si denominamos $\Theta = (\theta_1^*, \theta_2^*, \dots, \theta_c^*)$ con $c = p + (M-1)q$, tendremos que los parámetros de la regresión se pueden calcular mediante la expresión:

$$\theta_i^* = \frac{\sum_{k=1}^c c^{ik} d_k}{|C|}$$

Siendo c^{ik} el adjunto del elemento (i, k) de la matriz C , y d_k el elemento k -ésimo de la matriz D .

2.3.2 Propiedades y características.

El estimador definido en el anterior apartado cumple propiedades de consistencia casi segura y de normalidad asintótica. Para eso vamos a suponer en lo que sigue que

S es un conjunto abierto en la topología usual de \mathbb{R}^p que contiene al soporte de f , (la función de densidad de X), que existe el momento de orden 4 de $\|X\|$, y que la sucesión

$$\{\delta_s: \mathbb{R}^p \times \mathbb{R}^p \rightarrow \mathbb{R}\}_{s=s(n) \rightarrow \infty}$$

verifica las siguientes condiciones:

- (2.1) a) $\sup_u \delta_s(x,u) = O(s^p) \forall x \in \mathbb{R}^p$
 b) $\delta_s(x,u) = \delta_s(u,x)$ y $\delta_s(x,u) \geq 0 \forall (x,u) \in \mathbb{R}^{2p}$
 c) $\int \delta_s(x,u) du = 1 \forall x \in \mathbb{R}^p$ con $\delta_s(x,u) = 0$ si $\|x-u\| > c\varepsilon(n)$ donde $c \in \mathbb{R}^+$ y $\varepsilon(n) \rightarrow 0$ cuando $n \rightarrow \infty$

Estas condiciones las cumplen por ejemplo el estimador **Histograma** y los estimadores **Kernel** con soporte compacto.

Para la obtención de los resultados de consistencia casi segura y normalidad asintótica fue necesaria la obtención previa de los siguientes resultados:

- (2.2) Proposición: Sea $g: \mathbb{R}^p \rightarrow \mathbb{R}$ una función lineal verificando $\|g(x)\| \leq \|x\|$, y sea $\{\delta_s: \mathbb{R}^p \times \mathbb{R}^p \rightarrow \mathbb{R}\}_{s=s(n) \rightarrow \infty}$ una sucesión de funciones medibles verificando las condiciones (2.1), entonces:

$$\int g(x) \delta_s(u,x) dx \rightarrow g(u) \text{ cuando } n \rightarrow \infty$$

(2.3) Lema: Si el cuarto momento de $\|X\|$ existe y son satisfechas las condiciones (2.1), se verifica:

1.- Si Y es acotada con probabilidad 1 y S es un conjunto compacto entonces:

$$\bar{V}_{ij}^r \xrightarrow{c.s.} V_{0,ij}^r, \quad \bar{V}_{ij}^s \xrightarrow{c.s.} V_{0,ij}^s, \quad \bar{T}_k^s \xrightarrow{c.s.} T_{0,k}^s, \quad \bar{T}_k \xrightarrow{c.s.} T_{0,k}.$$

$$\text{Donde: } V_{0,ij}^r = E[I^r x^i x^j], \quad V_{0,ij}^s = E[x^i x^j],$$

$$T_{0,k}^s = \sum_{h=1}^q \theta_h^s \cdot E[I^s x^k x^h] + \sum_{h=q+1}^p \theta_h \cdot E[I^s x^k x^h]$$

$$T_{0,k} = \sum_{h=1}^q \sum_{l=1}^M \theta_h \cdot E[I^l x^k x^h] + \sum_{h=q+1}^p \theta_h \cdot E[x^k x^h]$$

2.- Si $\sqrt{n} \left(E[\bar{V}_{ij}^r] - V_{0,ij}^r, E[\bar{V}_{ij}^s] - V_{0,ij}^s, E[\bar{T}_k^s] - T_{0,k}^s, E[\bar{T}_k] - T_{0,k} \right)' \rightarrow 0$ y existe $\gamma > 0$

tal que $E(|Y|^{2+\gamma}) < \infty$, entonces:

$$\sqrt{n} \left(\bar{V}_{ij}^r - V_{0,ij}^r, \bar{V}_{ij}^s - V_{0,ij}^s, \bar{T}_k^s - T_{0,k}^s, \bar{T}_k - T_{0,k} \right)' \xrightarrow{d} N_c(0, \mathcal{A})$$

siendo \mathcal{A} una matriz simétrica que tiene una fila y una columna por cada uno de los

elementos de la forma $\bar{V}_{ij}^r, \bar{V}_{ij}^s, \bar{T}_k^s, \bar{T}_k$ definiéndose los elementos a_{ij} de la matriz

en función de los elementos \bar{V} o \bar{T} a los que corresponda su fila y su columna:

- Elemento $(\bar{V}_{ij}^r, \bar{V}_{kl}^s)$:
 - Si $r \neq s$: 0
 - Si $r = s$: $Cov(I^r x^i x^j, I^r x^k x^l)$
- Elemento $(\bar{V}_{ij}^r, \bar{V}_{kl}^r)$: $Cov(I^r x^i x^j, x^k x^l)$

• Elemento $(\bar{V}_{ij}^r, \bar{T}_k^s)$: - Si $r \neq s$: 0

$$- \text{ Si } r = s: \sum_{h=1}^q \theta_h^r \cdot \text{Cov}(I^r x^i x^j, I^r x^k x^h) + \sum_{h=q+1}^p \theta_h \cdot \text{Cov}(I^r x^i x^j, I^r x^k x^h)$$

• Elemento $(\bar{V}_{ij}^r, \bar{T}_k)$: $\sum_{h=1}^q \theta_h^r \cdot \text{Cov}(I^r x^i x^j, x^k x^h) + \sum_{h=q+1}^p \theta_h \cdot \text{Cov}(I^r x^i x^j, x^k x^h)$

• Elemento $(\bar{V}_{ij}, \bar{V}_{kl})$: $\text{Cov}(x^i x^j, x^k x^l)$

• Elemento $(\bar{V}_{ij}, \bar{T}_k^s)$: $\sum_{h=1}^q \theta_h^s \cdot \text{Cov}(x^i x^j, I^s x^k x^h) + \sum_{h=q+1}^p \theta_h \cdot \text{Cov}(x^i x^j, I^s x^k x^h)$

• Elemento $(\bar{V}_{ij}, \bar{T}_k)$: $\sum_{h=1}^q \sum_{l=1}^M \theta_h \cdot \text{Cov}(I^l x^i x^j, x^k x^h) + \sum_{h=q+1}^p \theta_h \cdot \text{Cov}(x^i x^j, x^k x^h)$

• Elemento $(\bar{T}_i^r, \bar{T}_k^s)$: - Si $r \neq s$: 0

$$- \text{ Si } r = s: \sum_{v=1}^q \sum_{h=1}^q \theta_h^r \theta_v^r \cdot \text{Cov}(I^r x^i x^v, I^r x^k x^h) + \sum_{v=1}^q \sum_{h=q+1}^p \theta_h \theta_v^r \cdot \text{Cov}(I^r x^i x^v, I^r x^k x^h) + \\ + \sum_{v=q+1}^p \sum_{h=1}^q \theta_h^r \theta_v \cdot \text{Cov}(I^r x^i x^v, I^r x^k x^h) + \sum_{v=q+1}^p \sum_{h=q+1}^p \theta_h \theta_v \cdot \text{Cov}(I^r x^i x^v, I^r x^k x^h) + \sigma^2 E[I^r x^i x^j]$$

• Elemento (\bar{T}_i^r, \bar{T}_k) : $\sum_{v=1}^q \sum_{h=1}^q \theta_h^r \theta_v^r \cdot \text{Cov}(I^r x^i x^v, x^k x^h) + \sum_{v=1}^q \sum_{h=q+1}^p \theta_h \theta_v^r \cdot \text{Cov}(I^r x^i x^v, x^k x^h) +$

$$+ \sum_{v=q+1}^p \sum_{h=1}^q \theta_h^r \theta_v \cdot \text{Cov}(I^r x^i x^v, x^k x^h) + \sum_{v=q+1}^p \sum_{h=q+1}^p \theta_h \theta_v \cdot \text{Cov}(I^r x^i x^v, x^k x^h) + \sigma^2 E[I^r x^i x^j]$$

• Elemento (\bar{T}_i, \bar{T}_k) : $\sum_{v=1}^q \sum_{h=1}^q \sum_{l=1}^M \theta_h \theta_v \cdot \text{Cov}(I^l x^i x^v, x^k x^h) + \sum_{v=1}^q \sum_{h=q+1}^p \sum_{l=1}^M \theta_h \theta_v \cdot \text{Cov}(I^l x^i x^v, x^k x^h) +$

$$+ \sum_{v=q+1}^p \sum_{h=1}^q \sum_{l=1}^M \theta_h \theta_v \cdot \text{Cov}(I^l x^i x^v, x^k x^h) + \sum_{v=q+1}^p \sum_{h=q+1}^p \theta_h \theta_v \cdot \text{Cov}(x^i x^v, x^k x^h) + \sigma^2 E[x^i x^j]$$

El teorema que me garantiza las propiedades del estimador que estamos estudiando es el siguiente:

(2.4) Teorema: Con las mismas condiciones que el lema (2.3), se cumple:

1.- Si Y es acotada con probabilidad 1 y \bar{S} es un conjunto compacto entonces:

$$\hat{\Theta} \xrightarrow{c.s.} \Theta$$

2.- Si $\sqrt{n} \left(E[\bar{V}_{ij}^r] - V_{0,ij}^r, E[\bar{V}_{ij}] - V_{0,ij}, E[\bar{T}_k^s] - T_{0,k}^s, E[\bar{T}_k] - T_{0,k} \right)' \rightarrow 0$ y existe $\gamma > 0$

tal que $E(|Y|^{2+\gamma}) < \infty$, entonces:

$$\sqrt{n}(\hat{\Theta} - \Theta) \xrightarrow{d} N_c(0, \mathcal{A})$$

siendo $\mathcal{A} = \sigma^2 \left(E[A(X)A'(X)] \right)^{-1}$.

3. LA OFERTA DE TRABAJO EN LAS COMUNIDADES AUTÓNOMAS.

La aparición del estado de las autonomías ha traído consigo el incremento del interés por el estudio de la situación de las comunidades autónomas y, como consecuencia, la preocupación por recoger más y mejores estadísticas con las que realizar mejores análisis de esa situación.

Se encuentran así, con mayor frecuencia, grupos homogéneos de datos sobre todas las comunidades, los cuales conforman paneles de datos con 17 submuestras (una por cada comunidad) que se prolongan a través del tiempo. Para obtener resultados más eficientes se hace necesario emplear técnicas de datos de panel.

El hecho de haber sugerido una nueva metodología de estimación de modelos haciendo uso de datos de panel y que este tipo de datos es el que se presenta en las comunidades autónomas, sugirió hacer una aplicación de dichas técnicas a la oferta de trabajo en las comunidades autónomas. Para ello se trataría de analizar si el comportamiento de las distintas comunidades es homogéneo respecto a la oferta de trabajo y cuales son algunos de los principales determinantes de ésta. Previamente se da una visión de la evolución de la "cuestión autonómica" desde su planteamiento moderno en el siglo XIX hasta la situación actual.

3.1. LAS COMUNIDADES AUTÓNOMAS EN ESPAÑA.

3.1.1. El estado de las autonomías.

En el marco de la Constitución de 1978 aparece el estado de las autonomías como una forma descentralizada de administración y un intento de dar respuesta a las sensibilidades nacionalistas y regionalistas que existían en diversas partes de España desde el siglo pasado (Estévez (1985)). Esta cuestión era uno de los problemas pendientes en la creación del moderno estado español y motivo de inestabilidad política durante varios periodos de este siglo.

Propuestas de estado con una cierta descentralización administrativa aparecen a lo largo del siglo XIX, siendo el primer intento de llevarlas a la práctica el proyecto de constitución federal de 1873, durante la I República. Sin embargo, ninguna se materializó hasta la constitución de la Mancomunidad Catalana, en 1914, creada al amparo del decreto de Mancomunidades de 18 de diciembre de 1913, el cual permitía la unión administrativa de varias provincias.

Al llegar la II República se diseña un nuevo modelo de estado en el que ya aparecen las autonomías como respuesta a las aspiraciones nacionalistas de Cataluña y País Vasco. En este modelo se reconoce la existencia dentro del estado de culturas particulares -Cataluña, País Vasco y Galicia - y se establece un modelo para el acceso a la autonomía regional de las nacionalidades que lo pretendan; sin embargo, parecía mostrarse reticente con el acceso a la autonomía de otras regiones (Fusi, (1989)). Este proceso se cortó con la guerra civil y no es hasta la restauración de la democracia cuando se vuelve sobre el tema.

Con la democracia, en la constitución de 1978, se intenta atender los dos cometidos

que motivaron los anteriores intentos de descentralización: solución de los problemas vasco y catalán y descentralización administrativa (Sole (1985)), estableciendo, de hecho, dos vías de acceso a la autonomía: una "rápida" para las nacionalidades históricas (Cataluña, País Vasco y Galicia) y otra "lenta" para las demás. Se produce así en primer lugar el acceso a la autonomía de las comunidades históricas y de Andalucía, que decide en referéndum optar por la vía rápida. Posteriormente lo hacen las demás comunidades, en un proceso que remata en 1983, con la constitución de los parlamentos en las comunidades de "vía lenta", surgidos de las primeras elecciones autonómicas.

Esta nueva estructura territorial trajo consigo una mayor necesidad de conocimiento de las realidades autonómicas, necesidad que se vio incrementada por la entrada en la C.E.E. y la importancia que en ella tiene el conocimiento de las desigualdades regionales a la hora de decidir su política de inversiones.

Nos encontramos entonces con la creciente importancia que adquieren los estudios de la economía desagregada por comunidades autónomas de cara a una valoración objetiva de las desigualdades existentes entre ellas y al conocimiento de sus necesidades particulares. Así, al tener un mejor conocimiento de la realidad de cada una de ellas obtendremos unos criterios objetivos a la hora de elaborar una política que reduzca las desigualdades existentes, tanto entre las distintas comunidades como entre éstas y las demás regiones europeas.

El impulso que el Estado de las Autonomías dio a este nuevo enfoque ha provocado el aumento y la mejora de los paneles de datos sobre las autonomías y la necesidad de desarrollar y mejorar los métodos de estimación con este tipo de datos de cara a superar los problemas que pueden aparecer al trabajar en este nuevo contexto.

3.1.2 Situación económica en España y en las comunidades autónomas durante el periodo estudiado.

La economía española está fuertemente ligada a la economía del mundo occidental, y en particular a la europea. Sigue por lo tanto sus avatares, pero presentando dos particularidades:

La primera es la presencia de un cierto retardo, que hace que la evolución de la economía europea se reproduzca en España con un retraso de aproximadamente dos años.

La segunda es una mayor volatilidad, consistente en un coeficiente de elasticidad mayor que 1 (1,4 en Raymond (1990)) para la relación entre el crecimiento español y el europeo.

La evolución de la economía presenta entonces durante el periodo dos situaciones muy diferentes.

La primera dura hasta el año 1985. Son años en los que aún persiste el efecto de las crisis energéticas de los años 70, provocando una situación de bajo crecimiento del P.I.B. y de pérdida de empleo. A partir de 1986, la economía se recupera y comienza una etapa de expansión, en la que el crecimiento alcanza valores superiores a las demás economías del mundo occidental. Esta expansión comienza a ralentizarse en 1990 y seguirá disminuyendo en los años siguientes.

Respecto de la situación en las comunidades autónomas, sus economías son muy dependientes, tanto de la economía española como de la europea, si bien de esta última dependen generalmente a través de la economía española. El grado de esta relación no será, sin embargo, el mismo en todas ellas. Las características propias de cada una van a determinar distintas reacciones a las variaciones de la coyuntura internacional o española. En Raymond (1990) se clasifican las comunidades según el grado de la relación de su economía con la española o con la internacional. En esa clasificación, Baleares y Canarias aparecen como dos comunidades más ligadas a la situación europea que a la española, por causa de la importancia que el turismo tiene para sus economías. La mayoría de las comunidades, sin embargo, aparecen más influidas por la situación española que por la europea. Entre estas comunidades Raymond considera tres subgrupos, en función de su sensibilidad a las variaciones de esa situación.

Las más sensibles serían las comunidades más industrializadas: Madrid, Cataluña y País Vasco, en las que la tasa de crecimiento del P.I.B. presenta una elasticidad superior a 1 con respecto a la tasa de crecimiento del P.I.B. español. El grupo de comunidades en las que la elasticidad de la tasa de crecimiento de su P.I.B. con respecto a la tasa de crecimiento del P.I.B. español es cercana a 1 incluye Aragón, Navarra, Asturias, Castilla-La Mancha y Castilla y León. Comunidades menos sensibles a las variaciones en España son Andalucía, Cantabria, Galicia, Murcia y Valencia, en las que la tasa de crecimiento de su P.I.B. con respecto a la tasa de crecimiento del P.I.B. español tiene elasticidad inferior a 1.

Las dos comunidades restantes, Extremadura y La Rioja, presentan una evolución de la situación relativamente independiente con respecto a la evolución de la economía

española, debida seguramente a características como su pequeño tamaño y el peso de la agricultura en su P.I.B., lo que provoca que factores autóctonos aumenten su importancia dentro de la variación del P.I.B. local a costa de la influencia de las variaciones de la situación española.

Para un análisis más particular de estas economías, Alcaide, Cuadrado, Fuentes (1990) agrupan a las comunidades en zonas con características comunes.

Los archipiélagos, Baleares y Canarias, poseen economías fuertemente basadas en el turismo, lo que las liga más a la coyuntura europea que a la española. La influencia de este factor provocó que la salida de la crisis se adelantase con respecto a las demás autonomías, (Aguiló (1990), De Lara (1993)). Así, a excepción del año 80 en el que repercute la segunda crisis del petróleo, los demás años hasta el 85 se produce un aceptable crecimiento que supera claramente al producido en el resto del estado. A partir del 85, la salida de España de la crisis se añade a su situación ya favorable y sus P.I.B. experimentan fuertes subidas, que colocan a Baleares en cabeza de las autonomías en P.I.B. por habitante y a Canarias cerca de la media nacional en esa misma magnitud.

La zona más dinámica en la península es la encuadrada en el arco mediterráneo: Cataluña, Murcia y Valencia. Favorecidas por su situación geográfica, estas autonomías experimentan durante la época de expansión unos crecimientos superiores a la media española, en cambio, durante la crisis, el comportamiento es

menos uniforme: Cataluña, con un fuerte peso de la industria en su economía (Carrau (1990)) sufre una recesión mayor que en las otras dos comunidades, Valencia y Murcia, en las que la mayor presencia del sector servicios suaviza los efectos de la crisis (Pedreño (1988)). Factores que influyen en la buena situación de estas economías son el modelo industrial de pequeñas empresas y producción de bienes finales, una actividad agraria moderna, basada en productos con grandes posibilidades exportadoras y la presencia de un sector turístico importante.

La otra zona expansiva de la economía española es el Valle del Ebro, que ocupa las comarcas de la ribera del Ebro a través de las comunidades de La Rioja, Navarra y Aragón, si bien la mayoría de las demás comarcas de estas comunidades no se ven afectadas por esta expansión.

Esta zona posee una situación muy ventajosa, fronteriza con Francia y dentro del triángulo marcado por el arco mediterráneo, Madrid y Bilbao. Además en su economía tiene fuerte presencia una actividad industrial diversificada y son comunidades con un P.I.B. por habitante alto entre las comunidades españolas. Estos factores contribuyeron a que la crisis fuese bastante suave para estas comunidades (ayudados en Aragón por la instalación de la General Motors (Serrano (1990)). Más tarde, durante la época de expansión, se refuerza esta ventaja, convirtiéndose en la segunda zona expansiva de la península (Alcaide, Cuadrado, Fuentes (1990), Alcaide (1992)).

Un factor primordial en la economía de Madrid fue desde siempre su capitalidad,

debida a la cual el sector servicios tiene un papel muy importante en su economía y existe una fuerte industria atraída por la cercanía de los centros de decisión y de la importante cantidad de población allí agrupada. La diversificación de su industria y la fuerza de los servicios amortiguaron el efecto de la crisis, viéndose menos afectada que la economía española general y que la mayoría de las autonomías. Sin embargo en la época de expansión su crecimiento pierde posiciones con respecto al crecimiento medio español. Factores como la descentralización administrativa de las autonomías o la apertura económica española que desplaza el centro económico del estado a comunidades más cercanas y mejor comunicadas con Europa influyen en este hecho. (García (1990)).

En el sur, Andalucía es una comunidad tradicionalmente situada entre las menos desarrolladas. La escasa presencia de industrias en declive hizo que la crisis no fuese más dura que la media española, mientras que, en la etapa expansiva, tuvo un crecimiento por encima de esa media, ayudada sobre todo por las fuertes inversiones públicas en infraestructuras, con motivo de los actos del 92 (Alcaide, Cuadrado, Fuentes (1990); Auriolés, Martín (1990); Auriolés (1992)).

En el norte, la Cornisa Cantábrica, formada por Asturias, Cantabria y País Vasco, presenta una situación de crisis permanente provocada por la masiva presencia en su economía de sectores en decadencia: minería, astilleros, siderurgia, ganadería, pesca. Debido a eso su crecimiento es débil incluso durante el periodo de expansión, y sólo en el año 1989 parecen unirse a la expansión estatal, para volver a caer en el noventa,

con el comienzo de la nueva crisis (Villaverde (1990); Cuervo, Vázquez (1990), Velasco, Díez, García (1990); Alcaide (1992)).

Las comunidades del interior, Castilla y León, Castilla-La Mancha y Extremadura, están tradicionalmente entre las comunidades menos desarrolladas. Están muy ligadas a una agricultura tradicional, si bien comienzan a aparecer núcleos de industrialización, en el triángulo Burgos, Palencia, Valladolid para Castilla y León y en los ejes Madrid-Guadalajara y Madrid-Toledo para Castilla-La Mancha (Alcaide, Cuadrado, Fuentes (1990)). En Extremadura tiene una presencia importante el sector energético, sin embargo, la práctica inexistencia de industrialización provocó que su economía se distanciase más de la media española durante la expansión (Pampillón, Vega, Espárrago (1990)).

Para finalizar, Galicia. Tiene una economía difícil de enmarcar en el esquema expansión-retroceso de la economía española (Alcaide, Cuadrado, Fuentes (1990)). Debido a una dependencia grande de la agricultura y la pesca tradicionales, Galicia soporta mejor que la mayoría de las comunidades las etapas de crisis pero, durante la expansión, la fuerte presencia de sectores en declive (pesca, astilleros, ganadería) hace que pierda la ventaja adquirida y su ritmo de crecimiento sea inferior que el de la economía española (Quintás (1990)).

3.2 DEFINICIÓN DEL MODELO DE LA OFERTA DE TRABAJO.

La enorme tasa de desempleo que afecta a los países desarrollados y sus graves consecuencias sociales son uno de los principales problemas que se le presentan a la economía en este fin de siglo. Su importancia se constata fácilmente en la cantidad de opiniones que sobre sus causas y sus remedios aparecen en todo tipo de publicaciones.

La necesidad de decidir cuales de estas ideas son realmente acertadas y en función de ellas encontrar una solución al problema provoca un creciente interés por la Economía Laboral.

Históricamente, las conjeturas sobre si variaciones en el salario provocan variaciones en el interés por trabajar o en el esfuerzo de trabajo pueden ser seguidas hasta los mercantilistas¹. Sin embargo, trabajos más cuidadosos sobre el tema son más recientes. Jevons (1888) cita ejemplos que le permiten deducir que la pendiente de la función de oferta laboral con respecto a los salarios era negativa.

El primer gran esfuerzo empírico por examinar la pendiente de la curva laboral con respecto a los salarios fue hecho en "The Theory of Wages" (1934) de Paul Douglas. Después de esta publicación, hitos importantes en la investigación de la fuerza laboral fueron: Schoenberg, Douglas (1937); Woytinsky (1940); Durand (1948); y Long (1958).

Con respecto a las horas de trabajo, existen trabajos de Lewis (1957); Bry (1959);

¹ Se pueden ver referencias en Douglas (1934, pag. 270); Long (1958, pag. 40).

Jones (1961) y Finegan (1962).

La investigación moderna en oferta laboral está caracterizada por un intento más cuidadoso de separar la medida de los efectos renta y sustitución. Data del artículo de Mincer (1962) el estudio sobre la participación en la fuerza laboral de las mujeres casadas y de la tesis de Kesters (1966) sobre las horas trabajadas por los hombres.

La aparición de los ordenadores a mediados de los 60 permitió la aplicación de técnicas estadísticas multivariantes a grandes cantidades de datos. Esta literatura fue objeto de numerosas panorámicas en trabajos como los siguientes: Heckman, Macurdy (1981); Heckman, Killingsworth, Macurdy (1981); Keely (1981) y Killingsworth (1981, 1983).

Otras publicaciones más recientes sobre oferta laboral son Pencavel (1986); Berkowitz, Anne Hill (1986); Blau, Ferber (1986); Lazear (1986); Willis (1987).

El modelo simple que se consideraba en el análisis tradicional del mercado de trabajo situaba al trabajador en la coyuntura de decidir, a cambio de un salario, si trabaja y cuánto tiempo hacerlo, de manera que se maximice su función de utilidad. Consideraba el tiempo del individuo dividido entre tiempo de ocio y tiempo de trabajo, con lo cual el ocio se convierte en un producto con un coste igual a la renta que se pierde por disfrutarlo. Con este enfoque se empezó a aplicar las nociones de efecto renta y efecto sustitución a la oferta de trabajo (Hicks (1932)).

En este modelo se consideraba que las decisiones del individuo (si trabajar y cuánto tiempo hacerlo) eran simultáneas, individuales, y sólo influía en ellas el momento actual. Modificando alguna de estas hipótesis se desarrollaron nuevos modelos sobre

la oferta de trabajo en los que las decisiones a tomar por el individuo no son simultáneas (Burdett, Mortensen (1978), Pissarides (1976)), o en las que se consideraban puntos de vista familiares o con perspectiva a largo plazo.

La evolución del modelo simple para situar al individuo en un entorno familiar comienza a aparecer en los años sesenta, con modelos en los que se consideraba el salario del esposo, en el modelo de oferta laboral de la mujer, como una renta no procedente del salario, mientras la oferta laboral del hombre era tratado de forma similar al modelo simple (Bowen, Finegan (1965, 1969)).

Un segundo modelo de oferta laboral familiar consideraba una utilidad familiar a maximizar, la cual dependía del consumo familiar total y del consumo de ocio de cada individuo de la familia (Samuelson (1956), Kusters (1966), Ashenfelter, Heckman (1974)).

El siguiente modelo aparecido, dejó de considerar la utilidad como algo conjunto para toda la familia y dio paso a utilidades individuales. Así la familia debe tomar una decisión sobre la oferta laboral y el gasto en consumo después de un proceso de negociación. (McElroy, Horney (1981)).

Otro enfoque muy empleado para el modelo de la oferta laboral es el de la producción familiar. Aquí se considera el tiempo dividido en tiempo de trabajo en el mercado y tiempo de trabajo fuera del mercado (el que se realiza para consumir bienes). Este tipo de análisis aparece con Becker (1965), pero tiene antecedentes en Mincer (1962, 1963). En Muth (1966) y Lancaster (1966) aparece este concepto más enfocado hacia el consumo. Becker (1974, 1975) amplía su trabajo inicial de un individuo al comportamiento de una familia. Trabajos críticos con este enfoque son

Barnett (1977), Pollack, Wachter (1974).

La perspectiva a largo plazo en la toma de decisión del individuo llevó a la aparición de modelos en los que se consideraban los efectos del retiro (Lazear (1986)) y, más importante, del tiempo dedicado a la formación y la aparición del concepto de capital humano (Becker (1965)).

Entre los temas tratados en los estudios sobre la oferta laboral tiene especial importancia la oferta laboral de la mujer, sobre la que se ha producido un gran número de trabajos (Finegan (1981), Lundberg (1981), Berliner (1989), Killingsworth, Heckman (1986), Meghir et al (1989)).

Publicaciones más completas sobre el mercado laboral son: Toharia (1983); Ashenfelter, Layard (1986); Elliot (1991); Kauffman (1991); Reynolds, Master, Moser (1991). Siguiendo a estos autores se presentan a continuación los principales elementos del mercado laboral, al igual que los factores que afectan a la oferta de trabajo para seleccionar de entre ellos aquellos más destacables en la economía española.

3.2.1 El modelo.

Como aplicación de las técnicas de datos de panel a datos producidos por la estructura de las comunidades autónomas españolas se va a estudiar un modelo de oferta laboral, y se contrastará la posibilidad de que presente diferencias entre las

distintas comunidades.

Si estas diferencias existen, su estudio puede proporcionar información que permita la mejora del modelo de partida. Además, si en la mayoría de las comunidades el modelo es homogéneo, se tendría una extensa fuente de datos con la que estimar un modelo global de la oferta laboral en España.

Las variables que se van a considerar en el modelo representarán el nivel de salarios, el nivel de educación y el nivel de riqueza económica de la comunidad.

Tradicionalmente se considera que las decisiones de ofertar trabajo o no y cuanto tiempo de trabajo ofertar son decisiones individuales de cada trabajador y uno de los factores fundamentales que tendrá en cuenta para tomar esa decisión será el salario que espera obtener a cambio (Reynolds, Masters, Moser (1991)).

Un salario mayor le permitirá comprar más bienes de consumo, por lo que estará interesado en incrementarlo lo más posible. Sin embargo, este incremento de salario puede traer consigo una disminución del tiempo de ocio, otro bien en el que también está interesado. Por esa causa, al tomar una decisión sobre su trabajo, el individuo intentará maximizar su satisfacción, buscando un equilibrio entre su tiempo de trabajo (que le proporciona un salario) y su tiempo de ocio. Aparece así la influencia del nivel de salario en la oferta laboral.

Sin embargo, esta influencia puede tomar dos formas: por una parte los salarios altos aumentan el atractivo del trabajo por lo que incrementarían la oferta laboral (efecto sustitución), por otra, como las necesidades de consumo son cubiertas más

rápidamente por salarios altos, también comienza más rápidamente a interesar el incremento del tiempo de ocio, esto provocara bajas en la oferta laboral (efecto renta). Las circunstancias económicas en las que se encuentra el individuo serán las que condicionarán cual de los dos efectos predominará.

Otra forma de relacionarse estas dos variables podría ser mediante la influencia de la oferta laboral sobre el nivel de salarios. La mayor o menor presencia de ofertantes de empleo actuaría sobre su "precio", que son los salarios, lo que los reduciría cuando la oferta sube y los incrementaría cuando baja.

La situación educativa de las CC.AA. afecta de manera fundamental a su creación de capital humano, por eso resulta de gran interés tener en cuenta el nivel de educación en el modelo (Kauffman (1991)).

La influencia de esta variable es debida a que, a medida que aumenta el nivel educativo, también acostumbra a aumentar el interés de los trabajos que se pueden conseguir, debido a factores como las mejores condiciones de trabajo o la consideración social que se tiene al realizarlos. Esto puede verse, por ejemplo, en la participación laboral que tiene la mujer en sociedades con mayor desarrollo cultural. En estas sociedades, el mayor de nivel de educación de las mujeres, hace que el no trabajar tenga para ellas un mayor coste de oportunidad, lo que las lleva a participar en el mercado laboral en mayor número que en las sociedades con menor nivel cultural. (Killingsworth, Heckman (1992)).

Por el contrario, la educación también produce un efecto renta, puesto que trae asociada una mayor valoración del tiempo libre y por lo tanto una reducción de la

oferta laboral. (Killingsworth, Heckman (1992))

Un tercer efecto se produce por la necesidad de los individuos de invertir un mayor número de años en su preparación en sociedades con más nivel educativo. Esto hace que se reduzca la oferta de trabajo con respecto a países con menor nivel de educación, donde los jóvenes se preparan menos y acceden antes al mercado de trabajo. En Bowen, Finegan (1965) encuentran una fuerte relación entre la tasa de actividad y el número de matriculados en las escuelas para los jóvenes entre 14 y 19 años. También obtienen resultados sobre ese aspecto Agüero, Olano (1988) y Novales (1988).

La tercera variable considerada va a ser el nivel de riqueza de la comunidad. A esta riqueza y a sus variaciones aparecen asociados diversos factores que pueden influir en la decisión de la oferta laboral.

Una comunidad rica puede ofrecer mejores niveles de protección social y de servicios públicos a su población, los cuales actuarían como un complemento de la renta de los individuos y podrían tener un efecto similar a una renta no procedente del trabajo (Kauffman (1991), Danzinger, Haveman, Plotnick (1981)).

También en relación con esta variable se tendría que en comunidades ricas existen mejores condiciones no salariales en los empleos, lo que haría más atractivo entrar en el mercado laboral (Killingsworth (1983)).

Las variaciones en la situación económica tienen asociadas además variaciones en la tasa de desempleo, la cual ejerce una influencia sobre la oferta de trabajo a través del *efecto del trabajador añadido* (una persona no activa se incorpora al mercado de

trabajo al quedar parado algún miembro de la unidad familiar) o del *efecto del trabajador desanimado* (los desempleados dejan de "buscar activamente" trabajo al desanimarse por no encontrarlo) (De Miguel, Lorente (1984), Tachibanaki, Sakurai (1991), Mincer (1966)).

Empleando esas tres variables se va a trabajar con el siguiente modelo:

$$O = \alpha_0 + \alpha_1 W + \alpha_2 E + \alpha_3 Y$$

Midiendo O la oferta de trabajo,

W el nivel de salarios,

E el nivel de educación,

Y el nivel económico de cada comunidad.

Los valores que se van a tomar para medir las distintas variables van a estar en función de la disponibilidad de sus datos para las comunidades autónomas. Se podría, en algún caso, buscar otros valores para estimar el modelo global pero, dado que el interés del trabajo está centrado en estudiar las posibilidades de los datos autonómicos, se ha preferido restringirlos a los disponibles para las autonomías.

Estos valores son:

Para medir la oferta laboral se usa la tasa de actividad, empleando datos publicados en "E.P.A. *Principales Resultados Trimestrales*".

Para medir el nivel de los salarios no se ha podido disponer de la ganancia media por hora de un trabajador por comunidades autónomas durante la mayoría de los años 80.

Debido a eso se empleará como medida el cociente entre la remuneración de asalariados y el número de asalariados de cada comunidad procedentes de las publicaciones "*Contabilidad Regional de España. Serie 1980-1987*" y "*C.R.E. Serie 1986-1990*", dividido por el I.P.C. de cada comunidad, con base en 1983, proporcionado por el I.N.E..

El nivel de educación se ha medido por la proporción del número de matriculados cada año en la enseñanza secundaria, procedentes de los anuarios estadísticos del I.N.E., con respecto al total de la población mayor de 16 años.

Por último, el nivel de la riqueza económica se ha medido mediante el valor añadido bruto por habitante, en pesetas constantes de 1983, obtenido en "*Contabilidad Regional de España. Serie 1980-1987*" y "*C.R.E. Serie 1986-1990*".

La importancia del efecto de los salarios en la oferta laboral es clásica en el análisis del comportamiento de la oferta y la demanda, sin embargo, la diversidad de comportamientos presentada en las Comunidades Autónomas puede dar pie al análisis de un efecto estructural en ambas variables. Para comprobar esta suposición se deben contrastar las diferencias entre los coeficientes estimados para cada una de las CC.AA. y el modelo de datos de panel clásico. Se hará una estimación autorregresiva en ambos casos. En el primero suponiendo un coeficiente de autocorrelación distinto para cada comunidad y estimado por el método de Prais-Winsten (Judge, (1988)) y el segundo considerando autocorrelación común a todas

las CC.AA., mediante datos de panel (Kmenta, (1986)).

Después mediante un test de especificación de Hausman (1978) se contrastará la diferencia entre los coeficientes estimados para el modelo conjunto y para cada una de las CC.AA.

El análisis clásico de datos de panel parte del hecho diferencial de aglutinar las variables de las distintas Comunidades, sin tener en cuenta las diferencias que se presentan en sus comportamientos a lo largo del tiempo, o bien en las interrelaciones espaciales.

Sin embargo, el hecho de trabajar con variables con tendencia produce efectos no queridos en las relaciones entre las variables, por ese motivo se analizará su comportamiento a largo y a corto plazo.

Estudiando las relaciones a largo plazo comprobaremos si las variables presentan comportamiento $I(1)$, lo que implicaría el peligro de que estas relaciones sean espúreas.

Por este motivo se analiza a continuación el comportamiento de cada una de las variables para decidir el modelo particular del que se hará uso.

Los resultados obtenidos en el estudio de este modelo se expondrán en el *Capítulo 4*.

3.2.2 Características de las variables.

Los datos empleados en este trabajo son datos trimestrales que cubren un periodo comprendido entre el 1º trimestre de 1980 y el 4º de 1990.

Oferta laboral.

Para medir la oferta laboral se empleó la tasa de actividad, o sea, el cociente entre el nº de personas activas y el nº de personas mayores de 16 años de cada comunidad.

Nivel salarial.

Se disponía de las series siguientes:

- *Remuneración de asalariados (serie anual) 1980-1990.*
- *Índices de ganancia media por hora trabajada en la industria y servicios para el total nacional (serie trimestral) I-81, III-89.*

Con estos datos se decidió emplear para medir el nivel salarial el cociente entre la remuneración de asalariados y la cantidad de empleo asalariado en cada comunidad autónoma.

Para trimestralizar estas cantidades se completó primero la serie de los índices entre I-80 y IV-90, y después se siguió el proceso que se indica en el *Apéndice III.1.*

Los valores así obtenidos se dividieron finalmente entre los I.P.C., con base 1983, de cada autonomía. Se tendría así unos valores reales para medir los salarios, que son los que se usará en el trabajo.

Nivel de educación.

Los datos con los que se contaba para medir este nivel son:

- *Número de alumnos matriculados en la enseñanza secundaria (B.U.P. y F.P., y, desde el curso 1986-87 también Reforma de enseñanzas medias) entre los cursos 1979-80 y 1990-91.*

No se poseían los datos de la Comunidad Valenciana del curso 1988-89, ni de Reforma en Galicia y Andalucía de ese mismo curso. Se realizó su estimación interpolando los valores para los dos años anteriores y los dos siguientes a las fechas que faltaban.

Además había algunos datos que parecen no ser correctos, ya que se apartaban claramente ^{de la} tendencia marcada por los demás². Esos datos son:

En Andalucía el curso 1986-87.

En Canarias los cursos 1982-83 y 1987-88.

En Extremadura los cursos 1985-86 y 1988-89.

En Murcia el curso 1986-87.

En el País Vasco el curso 1982-83.

Y en La Rioja el periodo comprendido entre los cursos 1979-80 y 1983-84.

Dados estos comportamientos extraños se decidió buscar algún tipo de suavización basada en las tendencias temporales deterministas más comunes (lineal y exponencial) y que mejor se adaptaban al comportamiento de los datos (ver *Apéndice III.1*).

Los datos así obtenidos se asignaron al cuarto trimestre del año en el que comenzaba el curso correspondiente.

² En todos estos casos el número de estudiantes cae súbitamente en los cursos mencionados, recuperándose en el siguiente. Rioja es una excepción puesto que la particularidad de los cursos mencionados es que son el único caso en el que durante un periodo desciende el número de estudiantes en una comunidad, por lo que se decidió su corrección.

Para los valores de los demás trimestres se procedió a la trimestralización de la serie, calculando la tasa media de variación trimestral:

$$\tau_n = \sqrt[4]{\frac{X_n}{X_{n-1}}}$$

y multiplicando: $X_{n-1} \cdot \tau_n^k = X_{n,k}$

Siendo X_n el número de matriculados en el curso que comienza el año n .

$X_{n,k}$ el número de matriculados que correspondería al trimestre k del año n .

Una vez trimestralizada la serie se dividieron los valores obtenidos entre los valores trimestrales de la población de mayores de 16 años de cada comunidad.

Para esta variable se hizo una excepción a la hora de calcular los valores para España. En lugar de repetir el mismo proceso que para los datos autonómicos, se decidió sumar los valores trimestrales de cada comunidad dividiendo después esta cantidad por la población española mayor de 16 años.

Nivel de riqueza.

Disponemos para medir este nivel de los datos siguientes:

- *Valor añadido bruto a precios corrientes por años y comunidades autónomas.*

Para usarlos se procedió primero a su trimestralización, de manera similar a la realizada anteriormente para el número de matriculados en secundaria

Una vez trimestralizada la serie se procedió a expresarla en precios constantes de 1972, dividiendo sus valores entre el correspondiente IPC de la comunidad.

Para finalizar, se dividieron estos valores entre el nº de habitantes mayores de 16 años correspondiente, en cada momento y en cada comunidad.

3.2.3 Comportamiento a largo plazo de las variables.

En muchos trabajos clásicos se había puesto de relieve el peligro de existencia de relaciones espúreas entre variables que tienen tendencias estocásticas. A partir de los trabajos de Granger (1981) y especialmente de Engle, Granger (1987) se suministra un instrumento econométrico que ayuda a detectarlas. Para ello el primer paso es analizar si las variables presentan comportamiento $I(1)$, lo que implicaría que tienen tendencias estocásticas y, en consecuencia -salvo cointegración- sus relaciones a largo plazo son espurias.

Se contrastará esta posibilidad mediante el test de Dickey-Fuller ampliado, apareciendo sus resultados en el *Apéndice III.3*.

El resumen de los resultados obtenidos es el siguiente:

Nivel de salarios.

Para esta variable no se rechaza la hipótesis de que los datos sigan un modelo $I(1)$ y por lo tanto que los salarios sigan en España una evolución impredecible a largo plazo.

El comportamiento en las autonomías es similar, con la única excepción de Canarias,

donde la evolución siempre creciente de sus salarios, acentuada en la segunda mitad del periodo, provoca que el modelo aceptado sea $I(1)$ con constante.

Nivel de educación.

La evolución de esta variable en España fue un crecimiento constante y homogéneo durante todo el periodo, lo que se traduce en un modelo $I(1)$ con constante para los datos agregados españoles.

En las comunidades autónomas esta variable sigue un proceso de crecimiento similar, presentándose también modelo $I(1)$ con constante en Asturias, Baleares, Castilla-La Mancha, Valencia, Cataluña y País Vasco. En Madrid el modelo aceptado es $I(1)$ con tendencia y constante, por lo que la impresión es que en esta comunidad la tendencia esperada no es constante, sino creciente en el tiempo.

En las demás comunidades se presenta únicamente una tendencia estocástica, bien por que existan incrementos menores (Cantabria, Navarra, Extremadura y Rioja) o bien por incrementos fuertes pero con grandes variaciones a lo largo del periodo (Galicia, Andalucía, Canarias y Murcia). En estos casos el modelo aceptado es $I(1)$, sin tendencia ni constante.

Nivel de riqueza.

La evolución de esta variable en España, más o menos estable al principio y ascendiendo durante la segunda mitad del periodo estudiado, se traduce en la aceptación de un modelo $I(1)$ con constante. Sin embargo en la mayoría de las autonomías el modelo aceptado es $I(1)$ sin tendencia ni constante. Se acepta el

modelo I(1) con constante en tres comunidades con una evolución creciente durante la mayor parte del período: Madrid, Aragón y Baleares. Canarias, también presenta una evolución creciente durante la mayor parte del período pero más fuerte que las anteriores por lo que el modelo aceptado es I(1) con constante y tendencia. El mismo modelo que en Castilla-La Mancha, que no presenta un crecimiento tan fuerte, y en La Rioja, donde la evolución es opuesta a las demás comunidades (*gráfica III.3.17*), con un incremento inicial y una estabilización posterior que se convierte en descenso a medida que avanza el período.

Tasa de actividad.

La tasa de actividad sigue en España un comportamiento impredecible, aceptándose un modelo I(1) sin tendencia ni constante, lo mismo que en la práctica totalidad de las autonomías. Las excepciones son Asturias y La Rioja, donde la reducción que se produce en su tasa de actividad se traduce en la aceptación de un modelo I(1) con constante.

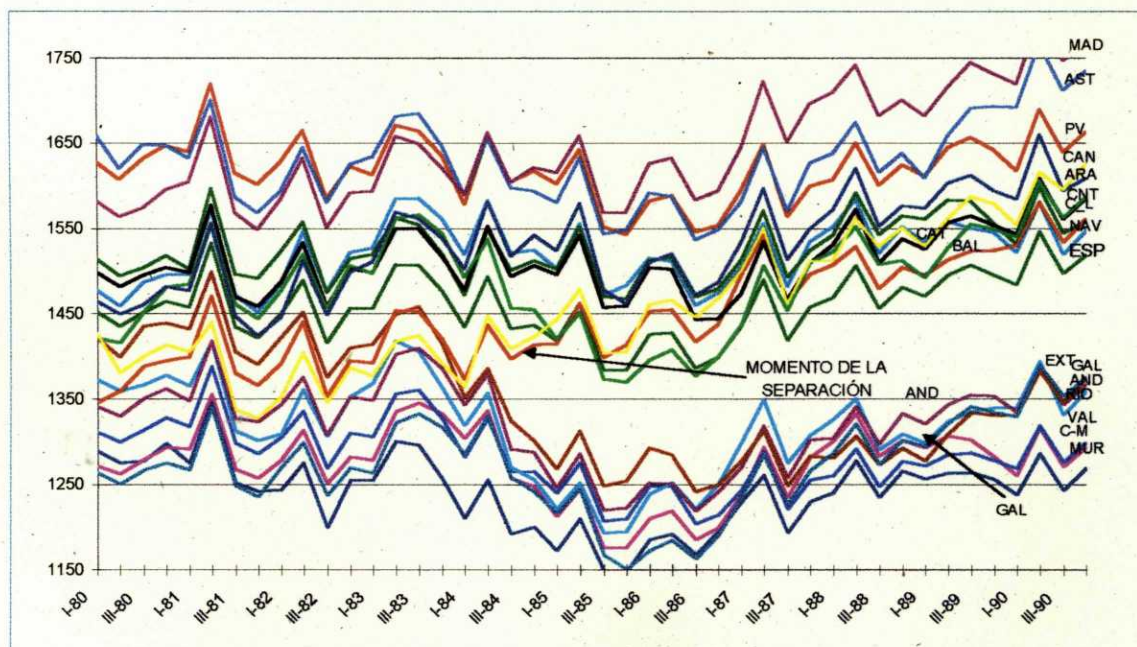
Se ha comprobado así que estas series presentan tendencia estocástica, por lo que se realizará el estudio con las primeras diferencias de las variables.

3.3 EVOLUCIÓN DE LAS VARIABLES QUE CARACTERIZAN LA OFERTA DE TRABAJO.

Nivel de salarios.

Lo primero que se puede observar en las gráficas del nivel de salarios (*Gráficas III.1.1 a III.1.18 en el Apéndice III.3*) es que presentan una cierta estacionalidad, con máximos en el segundo trimestre y mínimos en el tercero.

En general su evolución aparece influida por las circunstancias económicas, pudiéndose señalar en el tiempo estudiado tres periodos distintos. En el primero existe una cierta estabilidad, con oscilaciones en torno a un mismo nivel. Dura, dependiendo de la autonomía, hasta 1984 o 1985, años finales de la crisis en España. En la segunda etapa los salarios caen de forma generalizada. Sin embargo, lo hacen antes en las comunidades de salarios más bajos (en 1984) que en las de salarios más altos (en 1985), comenzando aquí un distanciamiento entre los dos grupos (*Gráfica 1*) que no se reducirá en la tercera

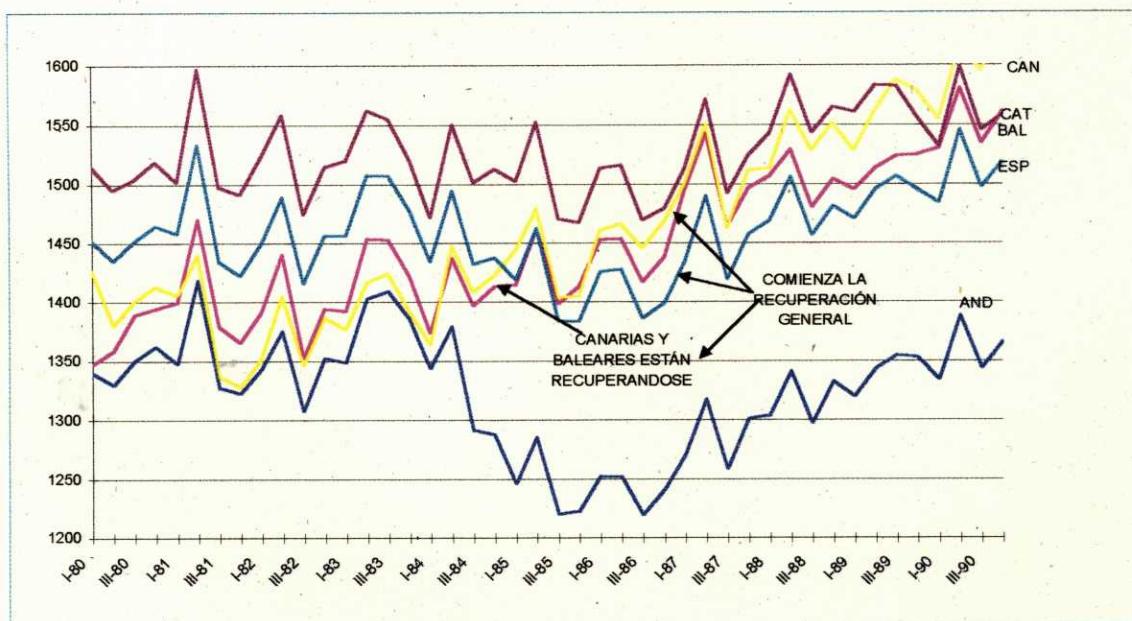


Gráfica 1

etapa, cuando, apoyados en la recuperación económica, los salarios inician una tendencia creciente que ya no abandonarán durante el resto del periodo.

Este esquema lo sigue también la evolución de la variable agregada para todo el estado (*Gráfica III.1.18*), pero con la peculiaridad de que aparece una suave caída en 1984, producida por los efectos combinados de las comunidades que comienzan a caer en ese año y de las que aún se mantienen.

Las dos comunidades en las que el comportamiento de esta variable se aparta más del general son Baleares y Canarias (*Gráficas III.1.4 y III.1.5*). La dependencia que el turismo causa en sus economías con respecto a la coyuntura europea (Raymond 1990) adelanta el crecimiento de sus salarios a 1985, dos años antes que lo habitual en las demás comunidades (la diferencia de tiempo entre la recuperación europea y la española) (*Gráfica 2*)

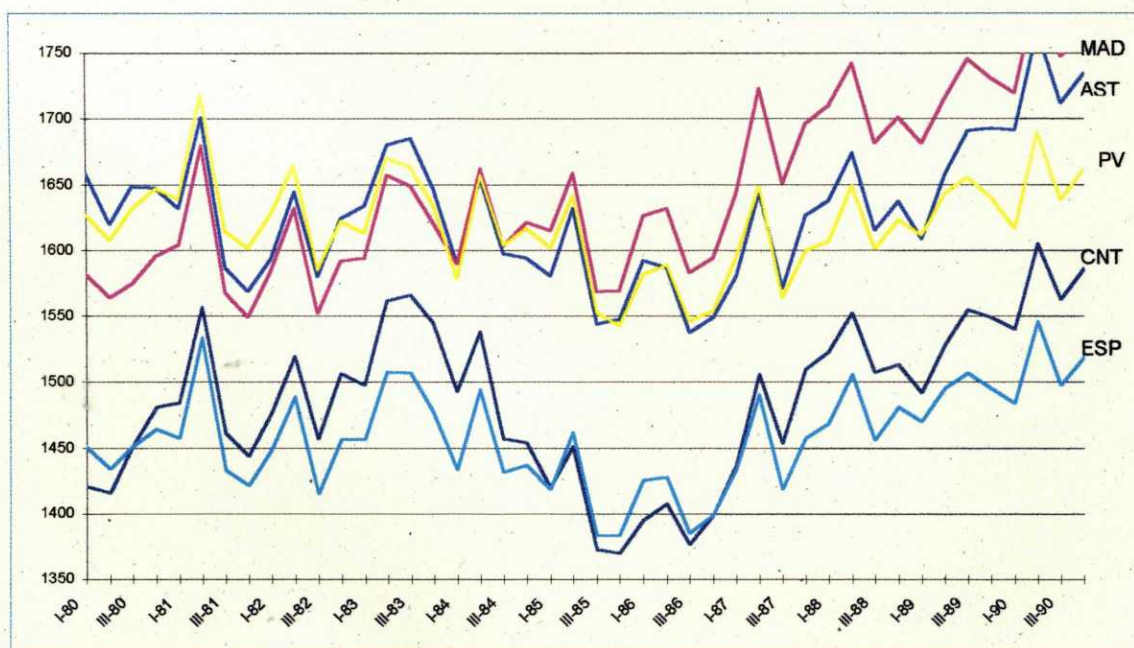


Gráfica 2

Las demás comunidades se pueden separar en dos grupos, según el año en el que se produce la caída de su nivel de salarios

En 1984 cae en Cantabria y en las comunidades de salarios más bajos (Andalucía, Castilla-La Mancha, Murcia, Galicia, Extremadura, Valencia y La Rioja), y en 1985 en las demás comunidades.

La caída de Cantabria al mismo tiempo que las comunidades de salarios más bajos es debida seguramente a la fuerte incidencia que la crisis tiene sobre las comunidades de la Cornisa Cantábrica, y que en las otras dos, Asturias y País Vasco, se manifiesta ya en la primera etapa, con una tendencia a la baja en lugar de la estabilidad observada mayoritariamente. Posteriormente, durante la etapa de crecimiento, los salarios en Asturias se recuperan con fuerza (influidos quizá por la fuerte presencia de empresas públicas en su economía) con lo que esta comunidad retorna a una

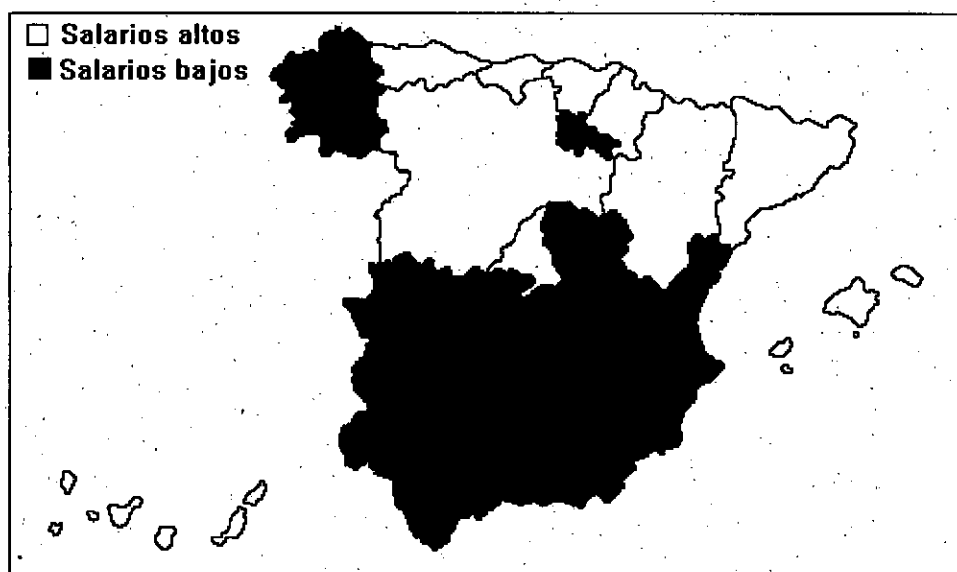


Gráfica 3

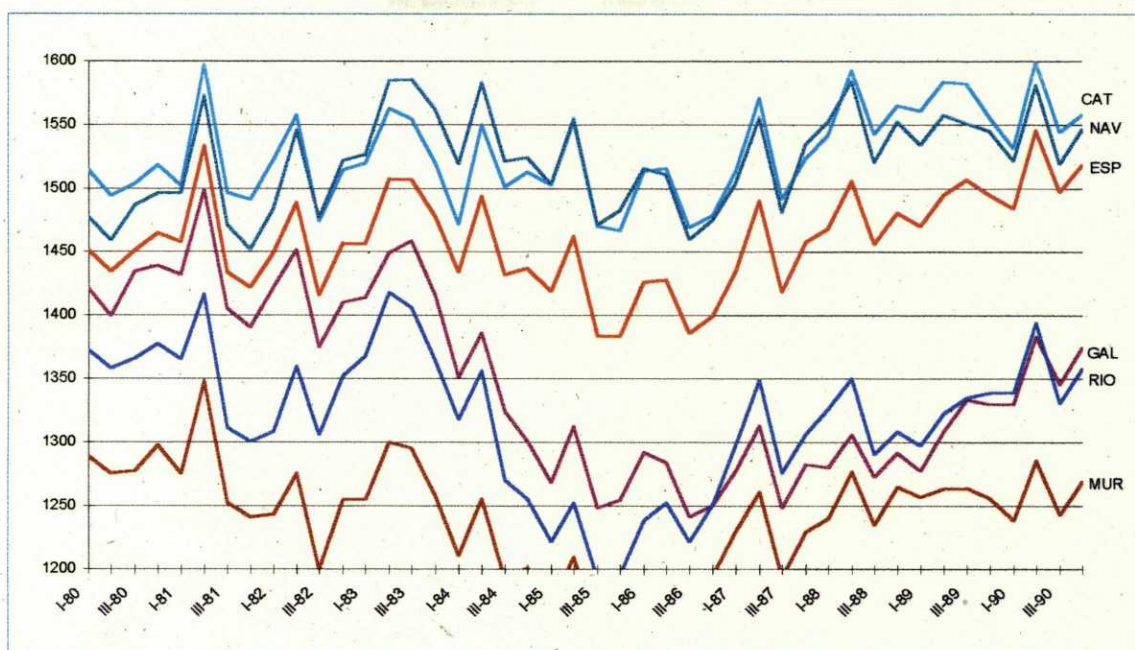
posición de cabeza, cercana a Madrid, mientras que Euskadi, que las acompañaba al principio del periodo, no consigue recuperarse. Con respecto a Cantabria, que es la primera de estas tres comunidades en comenzar su recuperación (*Gráfica 3*), también presenta una fuerte subida, que le permite volver al grupo de comunidades de salarios altos, del que la había apartado la fuerte caída sufrida en 1985.

El resto de las comunidades del grupo de salarios altos tiene un comportamiento más homogéneo (con la excepción de una subida final más acusada en Madrid), con un periodo estable inicial, caída durante un breve periodo (1986), y crecimiento posterior.

En el grupo de salarios bajos hay más heterogeneidad. Durante el periodo inicial, más que una estabilidad lo que se observa es una tendencia decreciente muy suave, que se agudiza al finalizar 1984, momento en el que tiene lugar una fuerte caída que



Mapa 1



Gráfica 4

abarca 1985 y 1986, y de la que salen en 1987, con un fuerte salto (excepto Valencia y Galicia que lo hacen de manera más suave), siguiendo hasta el fin del periodo con una suave tendencia creciente (excepto Extremadura en la que es más fuerte).

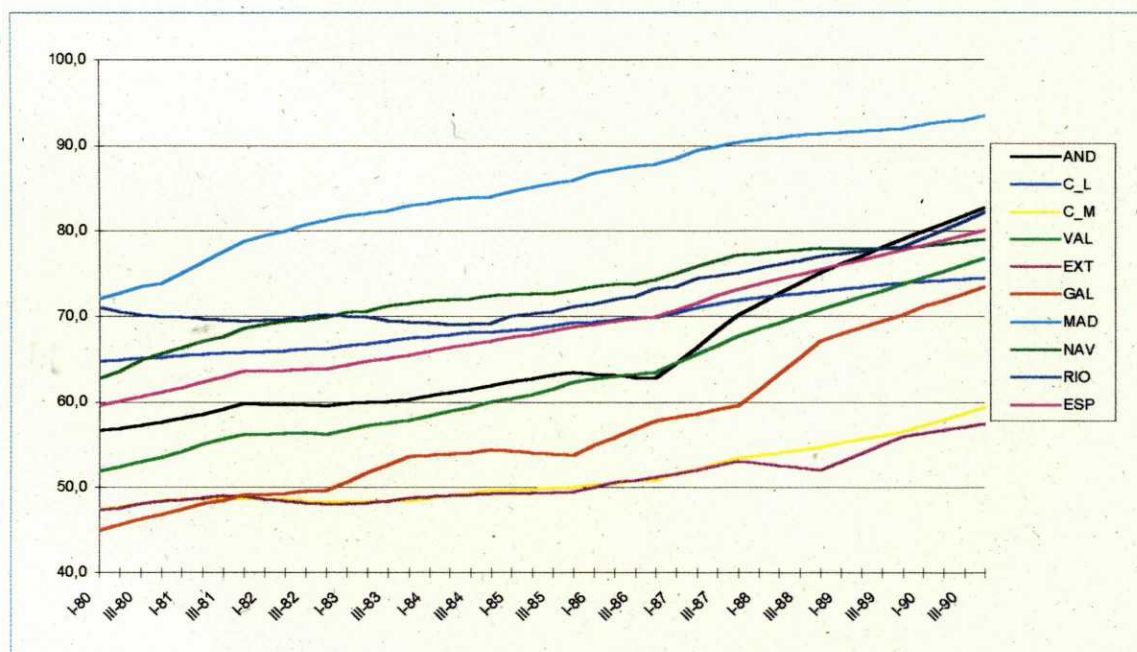
El reparto de comunidades entre estos dos grupos de salarios, altos y bajos aparece relacionado con su situación: Norte o Sur (*Mapa 1*). Las comunidades de salarios bajos tienden a ser las del sur, excepto dos: La Rioja, que podría estar afectada por la fuerte relación que sus sectores económicos tienen con la agricultura; y Galicia, con una economía poco desarrollada pero que, sin embargo, aparecía más cerca del grupo de los salarios altos que del grupo de los bajos al comienzo del periodo (*Gráfica 4*). Situación que se modifica por efecto de una crisis similar a la de la Cornisa Cantábrica (crisis láctea, pesquera, de astilleros), y a la que no regresa en lo que queda del periodo estudiado.

Nivel educativo.

Crece durante el periodo estudiado, con una pendiente constante que se incrementa a partir de 1987, tanto a nivel global como en las autonomías (*Gráficas III.2.1 a III.2.18*).

La evolución global (*Gráfica III.2.18*) presenta dos alteraciones en este crecimiento, en 1982 y en 1987. La primera coincide con una alteración presente en ese momento en la mayoría de las autonomías. La segunda es una alteración exclusiva de Andalucía (*Gráfica III.2.1*), que se traslada a la evolución global debido al tamaño de esta autonomía (vive en ella más del 20 % de la población española). La importancia del tamaño de Andalucía en la evolución global de este nivel también se observa en el aumento de la pendiente que ocurre a partir de 1987 y que es reflejo suavizado del ocurrido en Andalucía; con este incremento final, esta comunidad supera la media nacional a partir 1989.

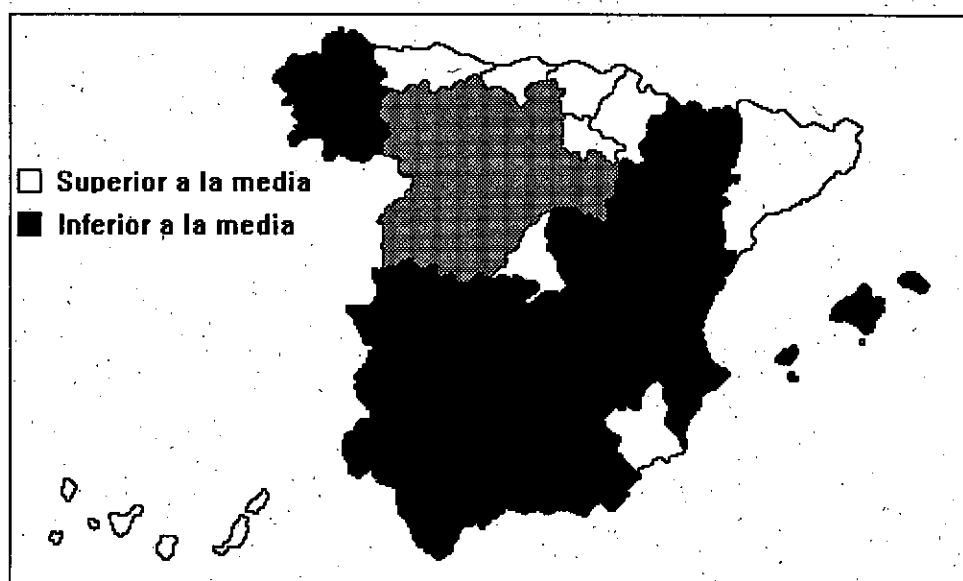
En las autonomías, la principal excepción a esta norma de crecimiento es La Rioja



Gráfica 5

(Gráfica III.2.17), que decrece o permanece estable hasta 1984. La causa de esta peculiaridad podría estar en su pequeña población, que la hace muy sensible a cualquier tipo de alteración. Coincide además esta peculiaridad con una etapa de incremento del P.I.B por habitante (Gráfica III.3.17) también exclusiva de esta comunidad.

La pendiente de las gráficas es muy semejante en todas las autonomías, siendo algo mayor en Galicia (que se acerca de forma acusada al nivel nacional) y Valencia; y menor en Extremadura (es la que más se va a separar del nivel nacional) y las dos Castillas (Gráfica 5). Observando los Cuadros 1 y 2, se podría tener un indicio que explicase este comportamiento en varias comunidades. Los cuadros presentan los incrementos relativos entre 1980 y 1990 de los valores promedios anuales para las variables salarios y nivel de educación. Se observa que Galicia y Valencia tienen grandes incrementos en nivel de educación y disminuyen en salarios. En Andalucía



VARIACIÓN DE SALARIOS	
Autonomía	Variación
CAN	13,73%
BAL	13,10%
MAD	11,31%
ARA	10,07%
CNT	9,08%
EXT	7,64%
C_L	5,32%
AST	5,01%
NAV	4,23%
CAT	3,34%
PV	1,43%
AND	0,92%
C_M	0,65%
RIO	-0,95%
VAL	-1,74%
MUR	-2,03%
GAL	-4,58%
ESP	4,22%

Cuadro 1

VARIACIÓN DEL NIVEL DE EDUCACIÓN		
Autonomía	Variación	Orden en salarios
GAL	57,31%	17
VAL	43,55%	15
AND	42,34%	12
BAL	36,13%	2
CAN	32,88%	1
AST	28,75%	7
MUR	28,33%	16
CAT	28,16%	10
ARA	27,66%	4
MAD	27,30%	3
CNT	25,87%	5
NAV	22,38%	9
C_L	21,83%	7
EXT	19,05%	6
PV	17,10%	11
RIO	14,56%	14
C_M	14,18%	13
ESP	31,21%	

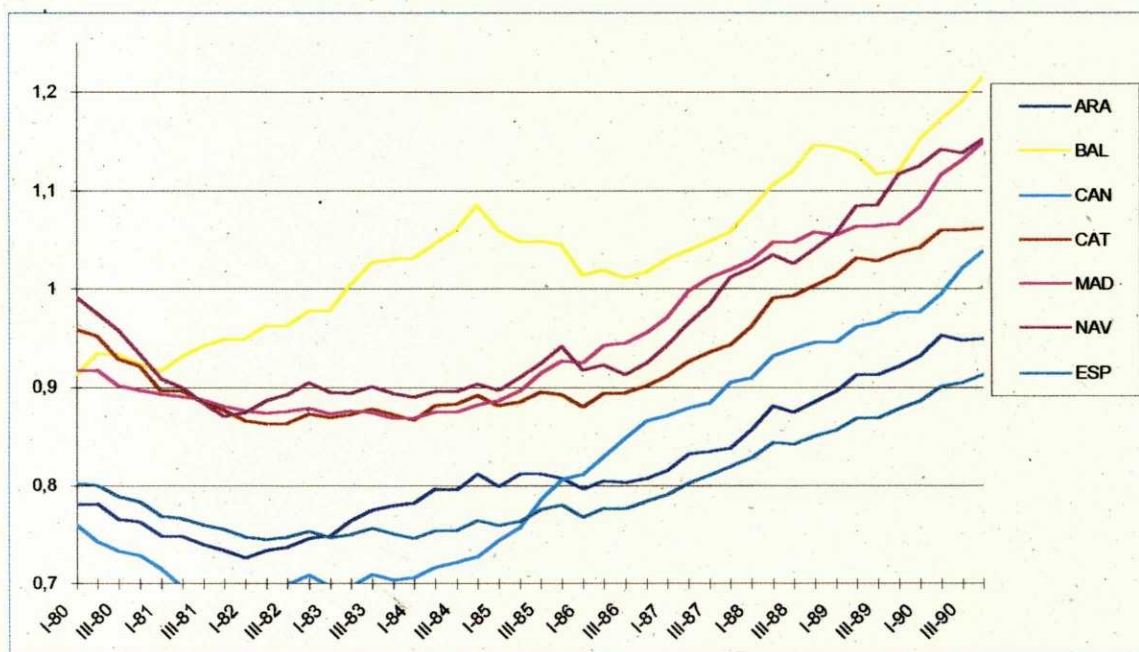
Cuadro 2

no disminuyen los salarios, sin embargo el incremento final es muy pequeño. Por la contra, Extremadura y Castilla y León han incrementado sus salarios a un nivel superior al de las comunidades pobres y están entre las que menos incrementan su nivel de educación.

Geográficamente (*Mapa 2*) las comunidades del sur (excepto Murcia) y Baleares aparecen con niveles por debajo del español, y las comunidades del norte (excepto Galicia) superan el nivel medio. Caso excepcional es el de Castilla y León, situada en un nivel cercano al español, pero encima en la primera mitad del periodo y debajo en la segunda

P.I.B. por habitante.

El perfil de esta variable para España es seguido más o menos fielmente y con más o



Gráfica 6

menos peculiaridades por las comunidades autónomas (*Gráficas III.3.1 a III.3.18*).

Empieza descendiendo hasta 1982, momento en el que se estabiliza con apenas un crecimiento muy suave (en las comunidades, este periodo oscila generalmente entre la estabilidad y un crecimiento suave). En 1986 (en algunas comunidades a principios de 1987), con la recuperación en marcha, se inicia un crecimiento que dura ya hasta el final del periodo estudiado. Este esquema es seguido bastante fielmente por Andalucía, Castilla-La Mancha, Castilla y León, Cataluña, Valencia, Murcia y Extremadura.

Comunidades que se comportan de una manera parecida, pero mostrando alguna particularidad son las siguientes:

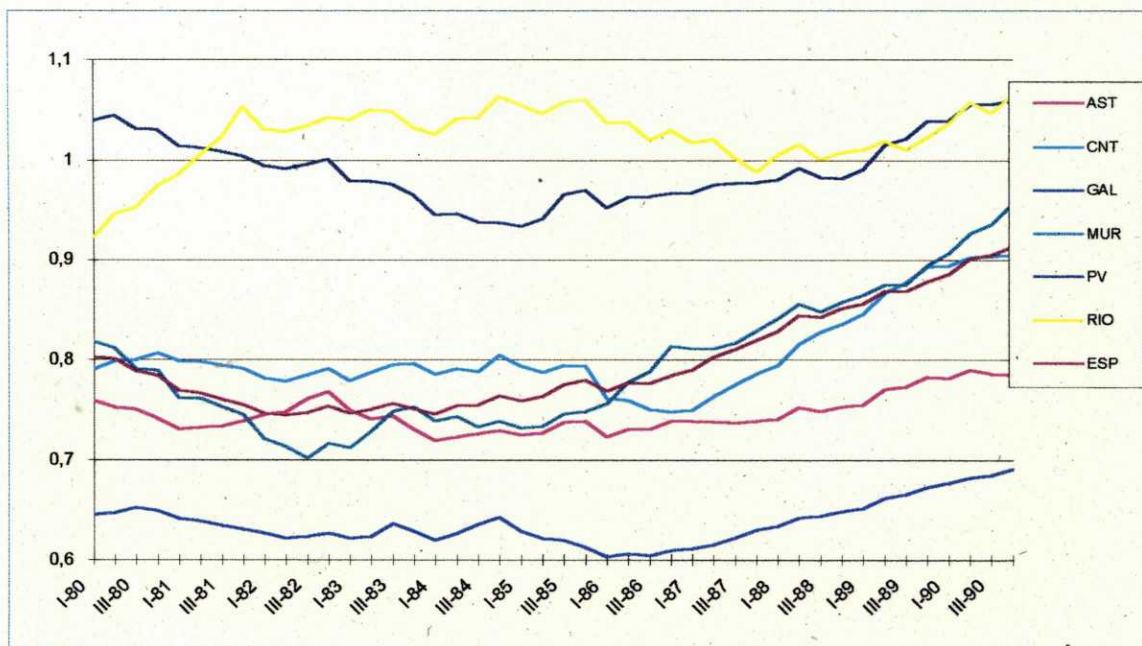
Aragón, que comienza a crecer ya en 1982, influido por la reactivación que en su

economía provoca la instalación de General Motors en Zaragoza (Serrano (1990)) (Gráfica 6).

Navarra, que presenta pendientes más acusadas, tanto en el periodo de caída como en la subida final. Estas pendientes son características de una economía industrializada, más sensible a crisis y expansiones que economías basadas en otros sectores. Esta característica, aunque menos acusada, también la presenta Cataluña. En Madrid también se observa una mayor pendiente, pero sólo en la etapa de crecimiento, que se adelanta además a la general. (Gráfica 6).

Galicia, que presenta pendientes más suaves. Causadas por la importancia que en su economía tenían la pesca y la agricultura tradicionales (Quintás (1990)), poco sensibles a las fluctuaciones de la economía. Esas pendientes suaves acercan su nivel al español durante la crisis y lo separan durante la recuperación. (Gráfica 7).

Las comunidades de la Cornisa Cantábrica (Gráfica 7). También presentan



Gráfica 7

pendientes más suaves que las generales. Otras particularidades suyas son:

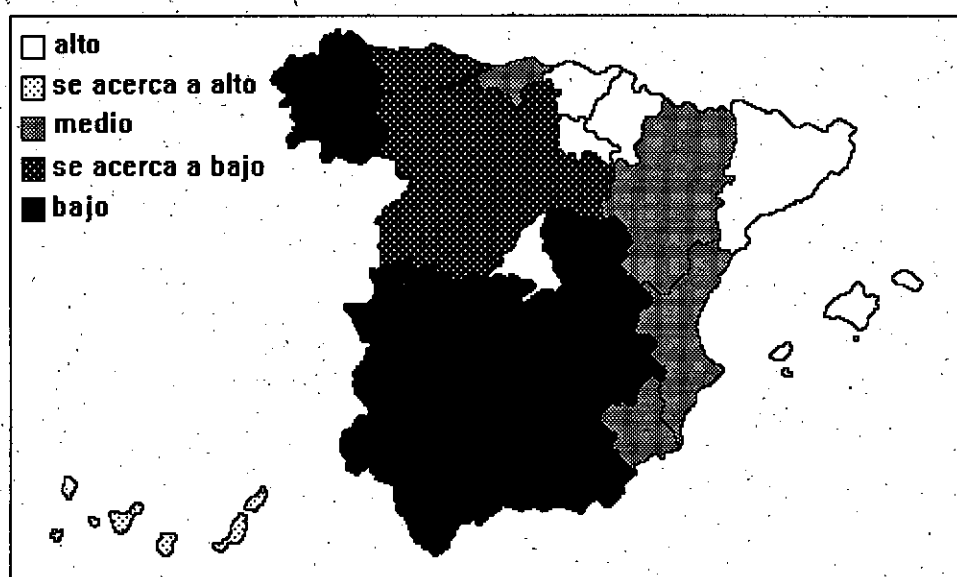
Euskadi, la más cercana al comportamiento general, mantiene la caída inicial hasta 1985 y su recuperación sólo es equiparable a la española a partir de 1989.

En Asturias la variable es prácticamente estable durante todo el periodo.

Cantabria también es estable hasta 1984, año en el que sufre una fuerte caída, de la que sale en 1987 con una fuerte recuperación.

Los archipiélagos (*Gráfica 6*), vuelven a ser comunidades con comportamientos muy diferentes al general, motivado por la fuerte relación de sus economías con el turismo, además el crecimiento de éste sector durante el periodo estudiado sitúa a Baleares en cabeza del P.I.B. por habitante de las comunidades españolas y acerca a Canarias en este aspecto al grupo de comunidades más desarrolladas.

Se observa sin embargo una diferencia entre ellas: Canarias cae hasta 1982, como la mayoría de las autonomías, y a partir de ese año sube continuamente hasta el fin del



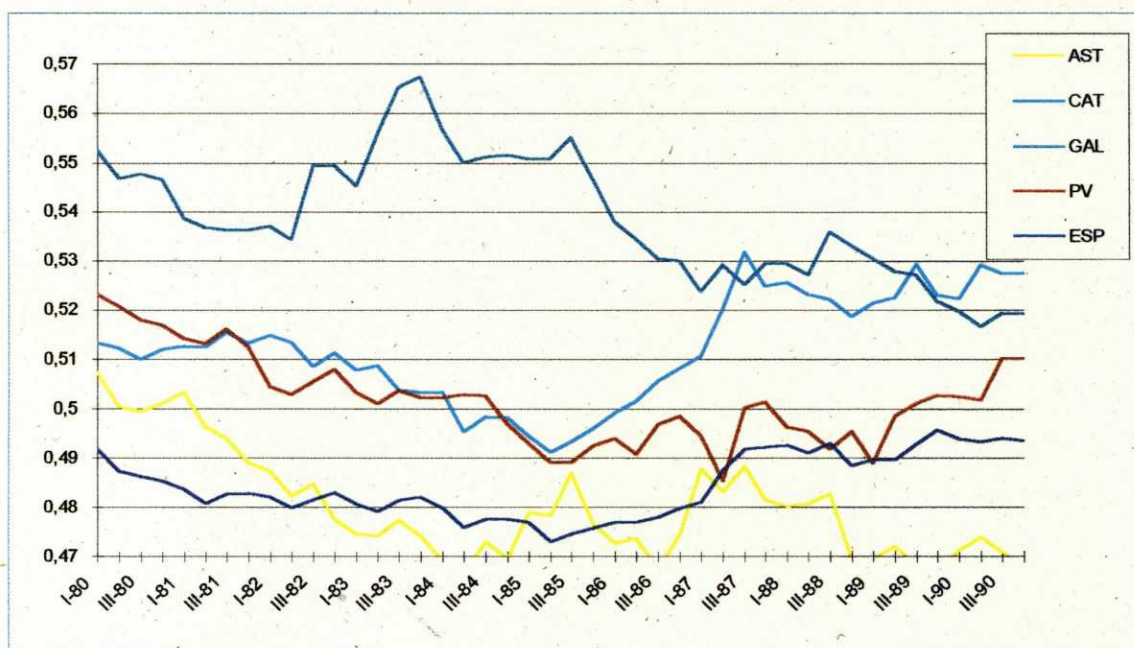
Mapa 3

periodo; en cambio Baleares no presenta esta caída inicial y sí otras dos, en 1985-86 y en 1989. El motivo de esta diferencia podría ser su distancia a los lugares de origen de los turistas, que convierte a Baleares en una opción preferente para el turismo en tiempo de crisis (desplazamientos más baratos), preferencia que es abandonada cuando la recuperación económica permite escoger opciones más caras.

El comportamiento más singular para esta variable es el de La Rioja (*Gráfica 7*), con un fuerte incremento en el periodo inicial (cuando las demás están bajando), estabilización entre 1982 y 1985, caída suave hasta 1988 (cuando las demás ya están metidas en la recuperación) y a partir de ahí un suave crecimiento hasta el fin del periodo. Una explicación para este comportamiento podría venir dada por su pequeña población y por la importancia que la agricultura y la transformación de sus productos tienen para su economía, por lo cual es fácil que se produzcan fuertes oscilaciones en su P.I.B. (Raymond (1990)).

Respecto de la posición que ocupan las autonomías en función de esta variable (*Mapa 3*), vuelven a aparecer abajo Galicia y comunidades del sur (Andalucía, Extremadura, y Castilla-La Mancha), con un acercamiento a esas posiciones de Asturias (comunidad en crisis) y Castilla y León. Arriba están Baleares y las comunidades industrializadas del norte: La Rioja, País Vasco (que también decae), Navarra, Cataluña y Madrid, acercándose Canarias a este grupo al final del periodo.

Las demás comunidades aparecen en una zona intermedia entre los dos grupos anteriores. Son tres comunidades de zonas en expansión (Aragón, Valencia y



Gráfica 8

Murcia) y Cantabria, a la que salva de caer de este grupo la rapidez (comparando con las otras comunidades de su zona) y la fuerza con las que inicia su recuperación.

Tasa de actividad.

El comportamiento general de esta variable es una tendencia descendente durante la crisis que se vuelve ascendente al comenzar la recuperación, en 1986. En varias comunidades y en la evolución global se observa un salto en 1987, en el momento en que se cambia la metodología de elaboración de la E.P.A. (Gráficas III.4.1 a III.4.18).

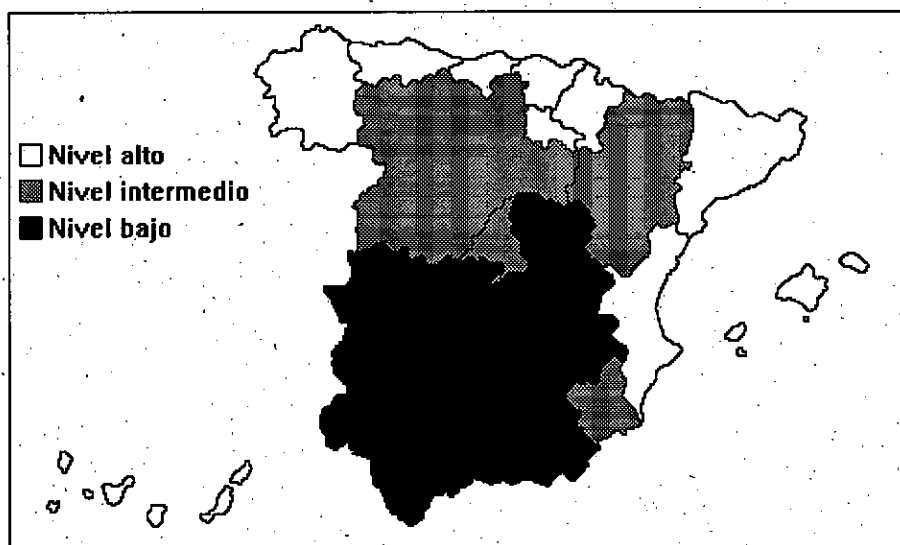
El hecho más destacable de esta variable es que Galicia sea la comunidad con tasa de actividad más alta durante la mayor parte del periodo (Gráfica 8), lo que contrasta con la tendencia general de que la tasa de actividad sea mayor en comunidades más desarrolladas. El origen de esta peculiaridad puede estar en la tradicional

participación de la mujer en labores fuera del hogar (agrícolas o marisqueo, por ejemplo), con lo cual la tasa de actividad femenina (en torno al 40%) es muy alta, incluso en comparación con las comunidades desarrolladas (Fernández-Jardon, Otero (1993) y Caneda, Otero (1994)). Al final del periodo se produce una caída en la tasa, al revés que la tendencia general. La combinación del envejecimiento de la población y del incremento del número de estudiantes (*Cuadro 2*) explicaría esta bajada. Las tasas de actividad de las demás comunidades, aunque también se ven afectada por estos dos factores, se compensan con el crecimiento de la tasa de actividad femenina.

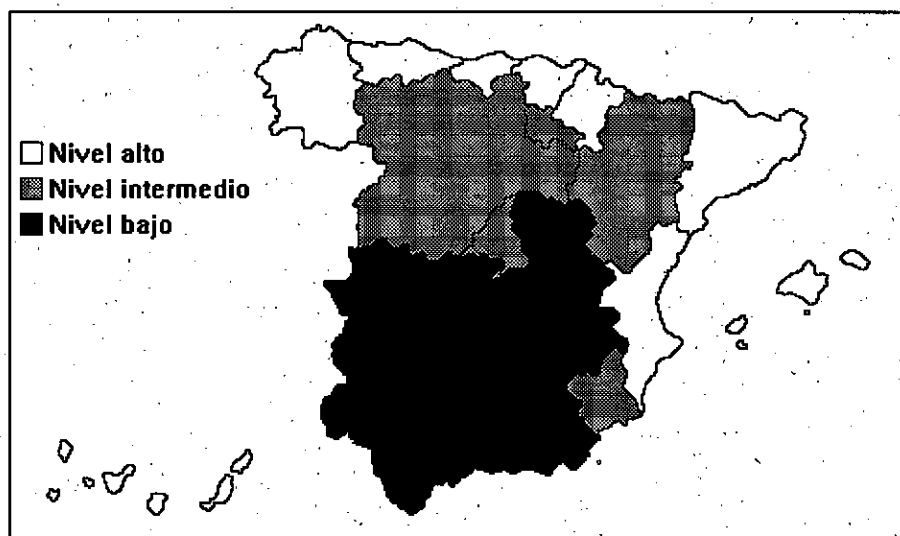
También es reseñable el hecho de que la tasa siga cayendo en Asturias durante todo el periodo, y que en el País Vasco sólo se recupere en 1989 (lo mismo que el P.I.B.) (*Gráfica 8*). Además, el tamaño de la población afecta a las oscilaciones de la variable. Así, las oscilaciones de España son menores que las observables generalmente en las comunidades, mientras que, excepto Navarra, las comunidades pequeñas (Baleares, La Rioja y Cantabria) presentan oscilaciones más acusadas que las grandes (*Gráfica 9*).

Geográficamente, esta variable aparece más ordenada que las anteriores, pero según dos esquemas: norte-sur e interior-exterior.

Para representar el comportamiento geográfico de esta variable vamos a usar las tasas medias anuales de los años 1980, 1985 y 1990 que aparecen en los *cuadros 3, 4 y 5* respectivamente.



Mapa 4



Mapa 5



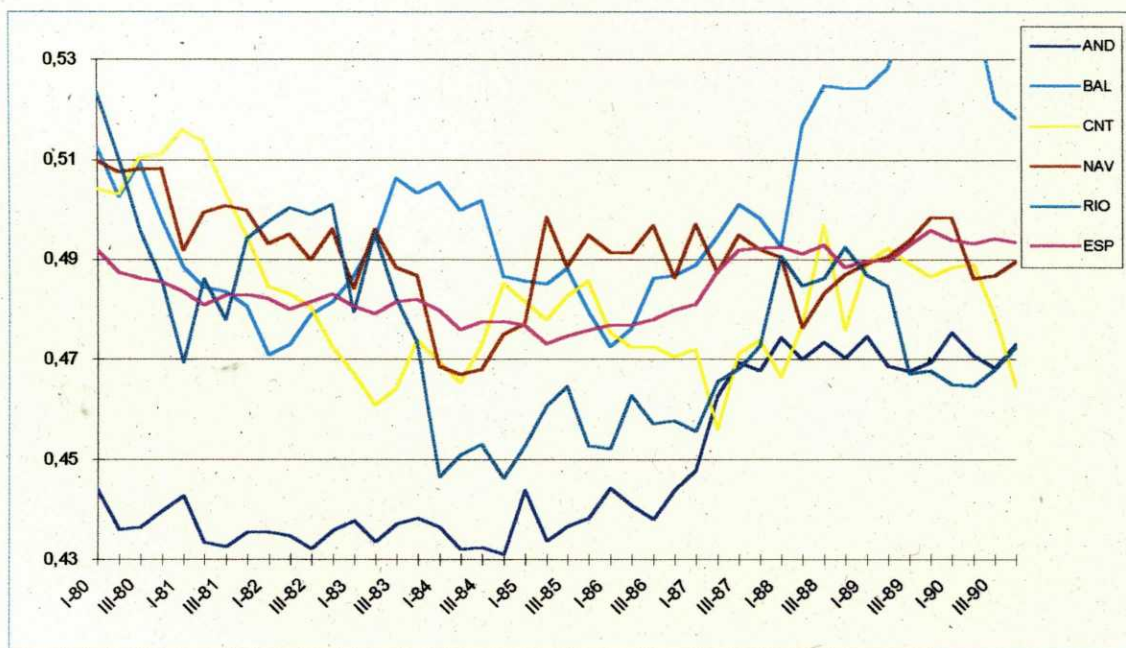
Mapa 6

Se representarán en tres *mapas* (n^{os} 4, 5 y 6) en los que se agruparán las comunidades según sus tasas medias, en tres niveles: alto, intermedio y bajo.

Se observa en ellos que la variable tiende a tomar valores altos en comunidades exteriores (las excepciones son Andalucía y Murcia, comunidades del sur), y disminuyen a medida que se avanza hacia comunidades interiores y del sur.

En 1980 y 1985 se conservan estas agrupaciones, pero en 1990 se observan algunas variaciones: las dos comunidades exteriores que no pertenecían al nivel alto ascienden de grupo (Andalucía al intermedio, Murcia al alto), y en el interior, Madrid toma unos valores que lo colocan a un nivel intermedio-alto, junto a dos comunidades que se descuelgan del nivel alto: Cantabria y Navarra. Del nivel alto también se descuelga Asturias, pasando al grupo del nivel intermedio.

Los cambios observados entre los dos últimos mapas parecen producirse en relación con variaciones del nivel económico de la comunidad. Asturias y Cantabria



Gráfica 9

Tasa de actividad promedio del año: 1980	
GAL	54.85%
PV	51.98%
CAT	51.18%
NAV	50.84%
CNT	50.72%
BAL	50.57%
CAN	50.54%
RIO	50.38%
AST	50.21%
VAL	49.72%
ESP	48.77%
ARA	48.44%
MAD	48.03%
C_L	47.85%
MUR	46.78%
EXT	44.48%
C_M	44.30%
AND	43.91%

Cuadro 3

Tasa de actividad promedio del año: 1985	
GAL	55.10%
CAN	50.05%
CAT	49.38%
PV	49.09%
NAV	48.98%
VAL	48.54%
BAL	48.47%
CNT	48.20%
AST	48.03%
ESP	47.51%
MAD	47.36%
MUR	46.61%
RIO	45.78%
ARA	45.63%
C_L	45.03%
AND	43.81%
C_M	43.58%
EXT	43.04%

Cuadro 4

Tasa de actividad promedio del año: 1990	
BAL	52.94%
CAT	52.67%
GAL	51.87%
VAL	51.23%
CAN	51.16%
MUR	50.76%
PV	50.62%
ESP	49.37%
NAV	49.02%
MAD	48.83%
CNT	48.02%
AND	47.19%
AST	47.09%
ARA	46.90%
C_L	46.89%
RIO	46.76%
EXT	45.16%
C_M	45.06%

Cuadro 5

pertenecen a una zona en crisis (el País Vasco también cae, si bien, no lo suficiente como para cambiar de nivel), Murcia (en una zona expansiva) y Madrid presentan una fuerte reactivación en la parte final del periodo (*Gráficas III.3.14 y III.3.13* respectivamente).

Un caso de difícil comentario es La Rioja, de hecho, en el estudio del comportamiento de la serie de la oferta de trabajo que se hace en el siguiente apartado, esta comunidad aparece con varios valores anómalos.

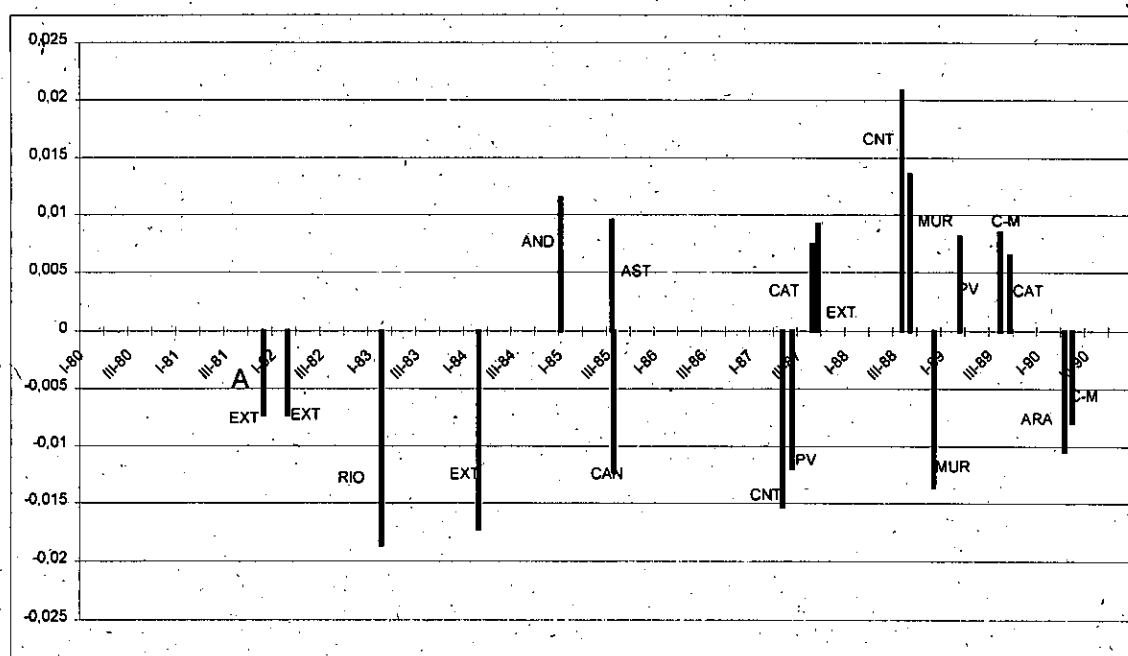
En el siguiente apartado, dentro del estudio del comportamiento de esta serie se comentarán con más detalle otras peculiaridades.

3.4 MODELIZACIÓN DE LAS SERIES DE LA OFERTA DE TRABAJO EN LAS COMUNIDADES AUTÓNOMAS.

Una vez vista la similitud -a nivel descriptivo- de la evolución de las series de la oferta laboral en las distintas comunidades, tiene interés, para comprobar las diferencias existentes, buscar el modelo que mejor se adapte al comportamiento de cada comunidad.

Las series de la oferta de trabajo están construidas con los valores trimestrales de la tasa de actividad de cada comunidad y de España proporcionados por la E.P.A., entre el primer trimestre de 1980 y el cuarto de 1990. Son por lo tanto 44 datos en cada comunidad y 44 datos para España.

Mediante diferentes test de raíces unitarias se comprueba que estas series presentan

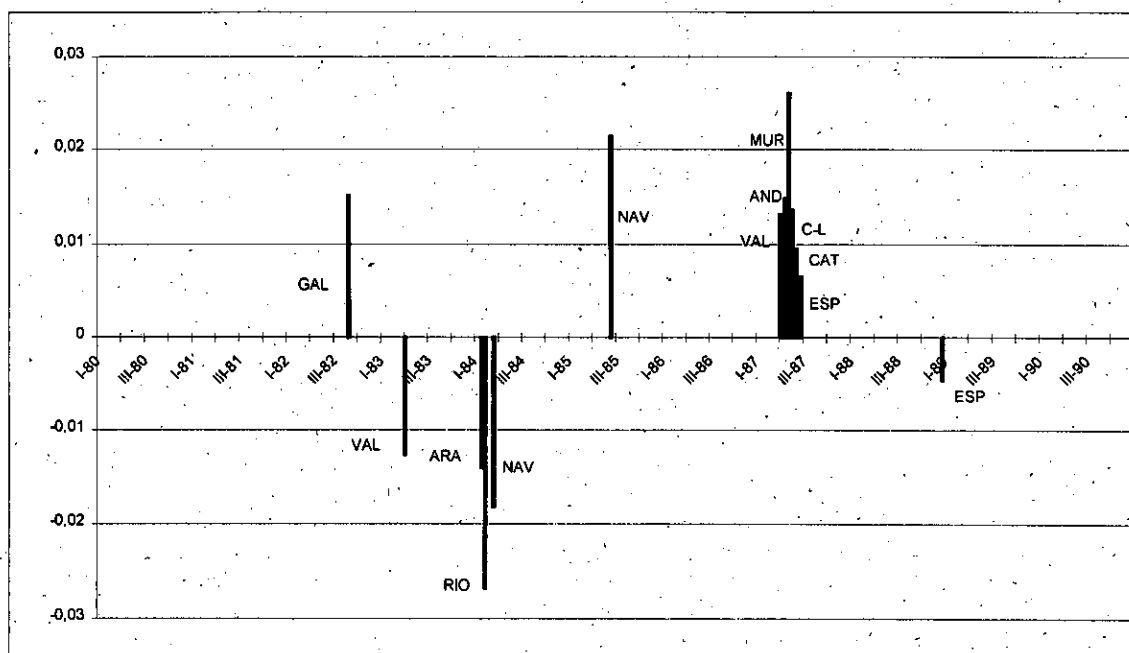


Gráfica 10: Impulsos

tendencia estocástica (*Apéndice III.4*) por lo que se realiza el estudio sobre sus primeras diferencias.

Los resultados de este estudio aparecen en el *Apéndice III.5*, y su resumen en los cuadros 6 y 7.

En una primera etapa se buscaron desviaciones que pudiesen interferir en el estudio del modelo teórico. Las desviaciones que se observaron parecen debidas a acontecimientos anómalos ligados en su mayoría a dos hechos: el cambio metodológico en la elaboración de la E.P.A. que comienza a aplicarse en el segundo trimestre de 1987 y el relanzamiento del empleo que se produce en algunas comunidades al salir de la crisis. Relacionados con el cambio metodológico se observan cinco escalones y dos impulsos precisamente en el trimestre en el que se produce (*Gráficas 10 y 11*) produciendo su acumulación un escalón también en

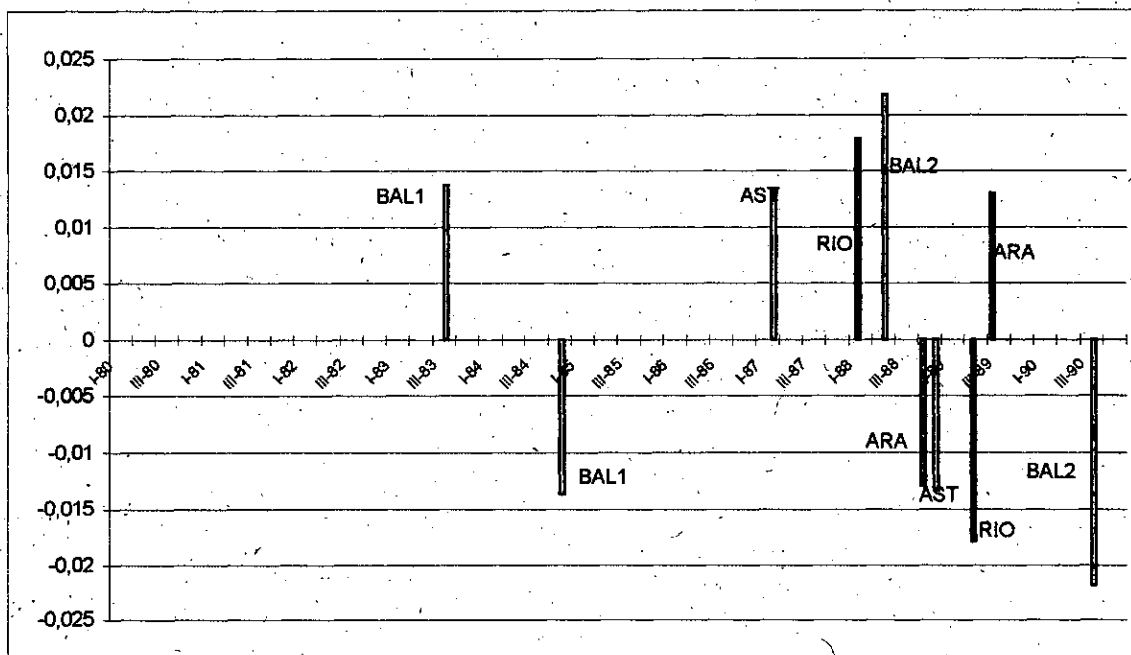


Gráfica 11: Escalones

España. Además el número de impulsos se incrementa a partir de ese momento, de forma que las dos terceras partes de ellos se observan en el tercio final del periodo estudiado (*Gráfica 10*)

Ligados a la recuperación económica parecen estar la mayoría de los escalones truncados (*Gráfica 12*) (excepto el que se produce en Aragón, que después de observar la gráfica de su tasa de actividad (*Gráfica III.4.2*) parece deberse más bien a un error metodológico). En Baleares, se produce el primero coincidiendo con la recuperación del turismo europeo (Europa está saliendo de la crisis en ese momento (Raymond (1990)) y el segundo con la recuperación del turismo español, (con la economía española ya metida en la etapa expansiva). En Asturias (Cuervo; Vázquez (1990)) y La Rioja (*Gráfica III.3.17*) la recuperación se produce más tarde que en las demás comunidades y parece que eso provoca un salto más intenso (si consideramos también el segundo escalón de Baleares se ve que aumentan su intensidad a medida que se alejan de 1985). La coincidencia en el cuarto trimestre de 1988 de dos escalones truncados (el final del de Asturias y el inicio del de Aragón) con un impulso en Murcia provoca el efecto de un escalón en la tasa de actividad española.

Otras desviaciones que merecen un comentario son los escalones que se producen en las comunidades del Valle del Ebro en el primer trimestre de 1984, las cuales, sin una causa clara, podrían ser debidas al efecto indirecto de las expectativas que se crean en la zona al instalarse la General Motors en Zaragoza, y que remitirían en ese momento a causa del mantenimiento de la crisis.



Gráfica 12: Escalones Truncados

Por lo que respecta al modelo teórico al que se ajustan las series de las diferencias de la oferta de trabajo en la mayoría de las comunidades se puede aceptar que se comportan como un ruido blanco:

$$\Delta X_t = \varepsilon_t$$

En consecuencia se tendrá que en éstas la tasa de actividad depende de si misma en el periodo anterior más una perturbación aleatoria simple; o sea, es totalmente imprevisible.

$$X_t = X_{t-1} + \varepsilon_t$$

Como excepción tenemos estacionalidad en seis comunidades: Andalucía, Aragón, Baleares, Canarias, Castilla-La Mancha y Castilla y León.

En el caso de Canarias y Baleares el peso que el turismo tiene en sus economías justifica la existencia de esa estacionalidad.

Aragón, Castilla - La Mancha y Castilla y León, son comunidades interiores en las que el clima provoca una fuerte estacionalidad en la agricultura, a la que se dedica un alto porcentaje de su mano de obra, que se vería afectada por esta estacionalidad.

En Andalucía podrían actuar los dos factores mencionados, sin embargo la estacionalidad de la agricultura puede estar mitigada por el P.E.R. con lo que tendríamos una estacionalidad producida más bien por el turismo. También presenta estacionalidad la variación de la tasa de actividad española. Aquí se refleja, sobre todo, el efecto del turismo en la economía, y en menor medida la estacionalidad de las labores agrícolas.

Además, al analizar el comportamiento a largo plazo de esta variable (*Apartado 3.2.3*) se aceptó la existencia de una constante para este modelo para dos comunidades: Asturias y La Rioja.

Existen otras cuatro comunidades autónomas en las que el comportamiento de las perturbaciones no es totalmente imprevisible, si no que recoge información de las perturbaciones anteriores. Éstas son Navarra, Valencia, Castilla-La Mancha y Cantabria.

Se analizó también la posible existencia de estacionalidad determinista en las comunidades (*Cuadro 7*). Los resultados lo rechazan, con una sola excepción: el País Vasco. Sin embargo al ser el único caso y no tener ninguna referencia que lo justifique, puede ser considerado como una consecuencia de la variabilidad de los resultados.

CC.AA.	Intervención				P. Estocástica		t-ratio media	desv. tip.
	tipo	fecha	coeficiente	t-ratio	Regular	Estacional		
ANDALUCÍA	Impulso	I-85	0.0115	4.07		AR(1):0.48	0.56	0.00390
	Escalón	II-87	0.0148	3.69				
ARAGÓN	Esc. trunc.	IV-88;III-89	-0.013	-4.03		AR(1):0.37	-0.28	0.00440
	Impulso	II-90	-0.0104	-3.24				
	Escalón	I-84	-0.014	-3.06				
ASTURIAS	Impulso	III-85	0.0095	3.25			-1.56	0.00391
	Esc. trunc.	I-87;IV-88	0.0133	4.58				
BALEARES	Esc. trunc.	II-88;III-90	0.0218	5.37		AR(1):0.37	-1.09	0.00560
	Esc. trunc.	III-83;IV-84	0.0137	3.36				
CANARIAS	Impulso	III-85	-0.0123	-2.83		AR(1):0.69	0.00	0.00608
CANTABRIA	Impulso	II-87	-0.0153	-3.35	MA(1):-0.42		-0.96	0.00625
	Impulso	III-88	0.0208	4.54				
CAST.-LM.	Impulso	III-89	0.0085	2.45	MA(1):0.42	AR(1):0.44	-0.11	0.00480
	Impulso	III-87	-0.0079	-2.26				
CAST.-LEÓN	Escalón	II-87	0.0136	3.34		AR(1):0.48	-0.93	0.00398
	Eliminado	IV-90						
CATALUÑA	Escalón	II-87	0.0095	3.27			-0.27	0.00280
	Impulso	III-87	0.0092	4.46				
	Impulso	III-89	0.0065	3.16				
VALENCIA	Escalón	II-83	-0.0126	-2.97	MA(1):-0.37		-0.35	0.00412
	Escalón	II-87	0.0131	3.09				
EXTREMADURA	Impulso	I-84	0.0073	-2.1			-0.17	0.00477
	Doble Imp.	I-82, IV-81	-0.0074	-2.99				
	Impulso	III-87	-0.0099	-2.82				
GALICIA	Escalón	III-82	0.0151	3.21			-1.64	0.00452
MADRID	Ningún						-0.12	0.00438
MURCIA	Escalón	II-87	0.026	4.22			-0.21	0.00601
	Imp. comp.	III-88;IV-88	0.0136	5.39				
NAVARRA	Escalón	I-84	-0.0182	-2.68	MA(1):0.54		-0.54	0.00659
	Escalón	II-85	0.0214	3.16				
PAIS VASCO	Impulso	II-87	-0.012	-5.09			-0.61	0.00325
	Impulso	I-89	0.0081	-3.42				
RIOJA	Impulso	I-83	-0.0187	-3.4			-0.50	0.00747
	Escalón	I-84	-0.0268	-3.45				
	Esc. trunc.	I-88;II-89	0.0179	3.26				
ESPANA	Eliminado	II-80				AR(1):0.57	0.78	0.00182
	Escalón	II-87	0.0065	3.51				
	Escalón	IV-88	-0.0046	-2.45				

Cuadro 6

CC.AA.	Residuos				Anómalos		ESTACIONALIDAD: t-ratio.					F-ratio
	r1	r2	r3	r4	fecha	valor/d.t.	d1	d2	d3	d4		
ANDALUCIA	-0.18 -0.18	-0.07 -0.11	0.00 -0.04	-0.05 -0.07	Ninguno supera 2.5 d.t.		1.88	-2.30	0.25	1.11	2.53	
ARAGON	-0.18 -0.18	-0.13 -0.17	0.06 0.00	0.06 -0.07	Ninguno		-1.81	0.36	0.82	0.01	1.02	
ASTURIAS	-0.23 -0.23	0.04 -0.02	-0.06 -0.06	0.01 -0.01	Ninguno		0.05	-1.53	0.11	-1.67	1.29	
BALEARES	0.11 0.02	0.02 0.01	0.11 0.10	-0.02 -0.05	Ninguno		-0.81	0.47	2.26	-1.61	2.15	
CANARIAS	0.00 0.00	-0.05 -0.05	-0.06 -0.06	-0.13 -0.13	Ninguno		1.35	-1.96	-0.26	0.41	1.48	
CANTABRIA	-0.09 -0.09	-0.13 -0.14	0.15 0.13	0.05 0.06	Ninguno		-1.04	0.00	-0.17	-0.32	0.30	
CAST.-LM.	0.05 0.05	-0.17 -0.17	0.12 0.14	-0.02 -0.07	Ninguno		-0.84	-1.15	0.54	1.56	1.18	
CAST.-LEON	-0.05 -0.05	-0.05 -0.05	-0.11 -0.12	-0.06 -0.07	Ninguno		-1.41	-0.46	1.95	-1.14	1.82	
CATALUNA	-0.05 -0.05	0.17 0.17	0.19 0.20	0.02 0.02	Obs ₁₈ :II-84	-2.79	0.38	-0.46	-0.01	0.63	0.19	
VALENCIA	-0.06 -0.06	-0.12 -0.12	-0.01 -0.03	-0.14 -0.16	Ninguno		-0.19	-0.88	0.26	1.45	0.74	
EXTREMADURA	-0.16 -0.16	0.31 0.29	-0.02 0.07	-0.01 -0.11	Obs ₄₃ :III-90	2.54	-0.18	-0.65	2.67	-1.48	2.44	
GALICIA	0.10 0.10	-0.11 -0.12	-0.15 -0.13	-0.06 -0.05	Ninguno		-3.19	-0.86	1.25	-0.81	3.29	
MADRID	-0.11 -0.11	0.23 0.22	-0.19 -0.15	-0.13 -0.21	Ninguno		0.26	-1.25	1.97	-0.23	0.96	
MURCIA	-0.01 -0.01	-0.16 -0.16	-0.07 -0.08	0.00 -0.03	Ninguno		-2.07	-0.16	1.66	0.89	1.96	
NAVARRA	-0.08 -0.08	0.09 0.08	-0.19 -0.18	0.05 0.02	Obs ₅ :I-81	2.92	-1.11	-1.50	-0.59	0.47	1.02	
					Obs ₃₄ :II-88	2.72						
PAIS VASCO	0.12 0.12	-0.20 -0.22	0.02 0.08	0.18 0.13	Obs ₄₃ :III-90	2.59	-2.74	-1.13	2.18	0.23	3.40	
RIOJA	0.05 0.05	0.14 0.14	-0.13 -0.14	0.03 0.02	Ninguno		-0.84	1.12	-1.19	-0.10	0.84	
ESPAÑA	0.13 0.13	0.02 0.00	-0.02 0.02	-0.11 -0.10	Ninguno		-0.76	-1.61	1.82	0.52	1.69	

Cuadro 7

4. ESTIMACIÓN DE LA OFERTA DE TRABAJO EN LAS COMUNIDADES AUTÓNOMAS.

En este capítulo se harán distintas estimaciones del modelo construido en el *Apartado 3.2* para comprobar las diferentes hipótesis planteadas.

Las dos primeras serán estimaciones globales, partiendo de los datos elaborados en los *Apartados 3.2.2 y 3.5*. Se empezará haciendo la estimación con los datos procedentes del conjunto de todas las autonomías y después se realizará con los datos agregados españoles.

Seguidamente se estimarán los parámetros del modelo para las distintas comunidades autónomas, y se contrastará en que comunidades este modelo difiere significativamente del global mediante un test de especificación de Hausman, buscando las causas que puedan explicar estas diferencias.

Posteriormente, eliminando del conjunto de datos los provenientes de las comunidades diferentes, se volverá a realizar la estimación global y se compararán estos nuevos resultados con la estimación inicial y con la obtenida a partir de los datos agregados.

El siguiente apartado estará dedicado a la estimación semiparamétrica.

En el se va a construir un estimador basado en el método descrito en este trabajo, empleándolo posteriormente para realizar las estimaciones global y por autonomías del modelo comparando sus resultados con los que se obtienen por mínimos cuadrados ordinarios.

4.1 ESTIMACIÓN MÍNIMO CUADRÁTICA DE LA OFERTA DE TRABAJO.

4.1.1 Estimación del modelo global.

Se calcula, en primer lugar, el efecto lineal de las variaciones de salarios, nivel educativo y nivel económico en la variación de la oferta laboral, según el modelo:

$$O = \alpha_0 + \alpha_1 W + \alpha_2 E + \alpha_3 Y \quad (\text{modelo 1})$$

Siendo O la oferta de trabajo,

W el nivel de salarios,

E el nivel de educación,

Y el nivel económico de cada comunidad.

El modelo será estimado inicialmente a partir de 731 datos obtenidos de las diferencias entre 44 trimestres en las 17 comunidades autónomas³.

La estimación se hizo mediante uso de datos de panel, donde se considera que existe heterocedasticidad y autocorrelación distinta en cada comunidad.

La primera hipótesis que nos interesa comprobar es si la variación de los salarios realmente afecta a la oferta de trabajo, o bien a la inversa, es la variación de la oferta la que condiciona la variación de los salarios. Para ello estimamos el modelo inicial

³ Los valores de los datos oscilaban entre un orden de 10^{-3} para O y 10^2 para W . Se homogeneizaron mediante los siguientes cambios de escala:

$$\begin{aligned} W' &= W; \\ E' &= 10^2 \cdot E; \\ Y' &= 10^3 \cdot Y; \\ O' &= 10^3 \cdot O; \end{aligned}$$

presuponiendo que la hipótesis nula es que la variación del salario no afecta positivamente a la oferta. Se ha optado por considerar sólo el signo positivo como alternativa porque el signo negativo implicaría un efecto renta, lo cual no parece factible, ya que los salarios y el nivel de vida en España durante la década de los 80 no fueron altos en comparación con los países de su entorno. Consecuentemente, la hipótesis que se va a contrastar es:

$$H_0: \alpha_3 \leq 0.$$

$$H_1: \alpha_3 > 0$$

Los resultados obtenidos son⁴:

$$O_t^* = -0.78401 + 0.0064687 \cdot E_t + 0.058184 \cdot Y_t - 0.012954 \cdot W_t$$

(0.1949) (0.002748) (0.01375) (0.003881)

En esta estimación resultan significativos todos los coeficientes estimados, si bien el de la variación del nivel educativo de una manera menos rotunda que los demás.

Sin embargo el coeficiente negativo del nivel de salarios no rechaza la hipótesis nula planteada anteriormente por lo que se optó por considerar el modelo alternativo, o sea, que son las variaciones de la tasa de actividad las que actúan sobre el nivel de salarios, (un incremento en la tasa de actividad provocaría un aumento en la oferta de trabajo, lo cual, si no existe un incremento de la demanda que lo compense, hará subir sus 'precios', o sea, los salarios) pasando así el nivel de salarios a ser variable

⁴ La salida completa puede observarse en el *Apéndice IV.2.*

explicada y la oferta de trabajo explicativa.

Entonces, a partir de los datos con los que se cuenta, se puede considerar más factible un modelo de la variación de los salarios en la que aparezcan como variables explicativas las variaciones de la tasa de actividad, del nivel económico (que actuaría sobre los salarios a través de la demanda de trabajo) y de la educación. Se tendría lo siguiente:

$$W = \beta_0 + \beta_1 O + \beta_2 E + \beta_3 Y \quad (\text{modelo 2})$$

Siendo W el nivel de salarios,

O la oferta de trabajo,

E el nivel de educación,

Y el nivel económico de cada comunidad.

Por consiguiente, las hipótesis a contrastar en dicho modelo son:

$$H_0: \beta_1 \geq 0.$$

$$H_1: \beta_1 < 0$$

Los resultados obtenidos son⁵:

$$W_t^* = - 3.1058 + 0.018525 \cdot E_t + 0.44941 \cdot Y_t - 0.47577 \cdot O_t$$

(1.473) (0.004735) (0.02721) (0.04901)

⁵ El análisis completo está en el *Apéndice IV.2*.

En este modelo siguen siendo significativos todos los coeficientes estimados, apareciendo en esta ocasión el término independiente como menos significativo que los demás.

Se obtienen para las variaciones de salarios relaciones directas con las variaciones de educación y nivel económico, e inversa con la variación de la oferta de trabajo. Consecuentemente se rechazaría la hipótesis nula planteada anteriormente, por lo que optaremos por suponer que la tasa de actividad actúa sobre el nivel de salarios por el lado de la oferta de trabajo: incrementos de la tasa de actividad significarán la entrada de gente en el mercado laboral y una mayor oferta de trabajo, con lo que se abaratarían sus precios. El efecto inverso se produciría cuando bajase la tasa de actividad.

La relación directa entre variaciones de nivel económico y de salarios se explicaría por la existencia de una relación directa entre crecimiento económico y demanda de trabajo: en épocas expansivas los empleadores demandan más trabajo lo que tenderá a elevar los salarios, mientras que esa demanda cae durante las recesiones, lo que tenderá a reducir los salarios.

El nivel de educación (proporción de estudiantes de secundaria) no sería en este caso una medida directa del capital humano, ya que sus cambios tardan un tiempo en reflejarse en el nivel global de la sociedad (se necesitan cuatro años para acabar los estudios). Sin embargo, su evolución en el periodo estudiado (crecimientos más o menos estables en todas las comunidades) serviría como un índice de la evolución del nivel de capital humano de las comunidades. Si se acepta esto, el coeficiente del

nivel de educación nos indicaría que el incremento de capital humano tiende a elevar los salarios, lo que coincide con lo esperado: mayor capital humano implicaría mayor cualificación y mayor rendimiento en el trabajo, lo que permitiría sostener mayores salarios.

A la vista de los resultados de la estimación de este modelo parece más lógico, para los datos con los que trabajamos, aceptar la hipótesis de que es el nivel de salarios el que está afectado por la oferta laboral y no al revés, por lo que se eliminará del modelo.⁶

El nuevo modelo quedará como sigue:

$$O = \theta_0 + \theta_1 E + \theta_2 Y \quad (\text{modelo 3})$$

Siendo O la oferta de trabajo,

E el nivel de educación,

Y el nivel económico de cada comunidad.

Estimándolo a partir de los datos iniciales el resultado obtenido es:

$$O_t^* = - 0.79218 + 0.0059438 \cdot E_t + 0.049101 \cdot Y_t$$

(0.2075) (0.002762) (0.01356)

⁶Otra posibilidad sería estudiar los retardos de la variable nivel de salarios, ya que se puede suponer que los efectos de la variación de los salarios no es inmediata, con lo que se podrían obtener resultados más aceptables con alguno de esos retardos. Sin embargo, como este trabajo está enfocado de cara al estudio de las posibilidades de los datos de panel y del manejo de datos que se abren con la estructura del estado de las autonomías, se ha preferido no seguir ese camino.

Los resultados obtenidos ahora para estas dos variables son parecidos a los obtenidos para el *modelo 1*: todos los coeficientes son significativos, pero el nivel de educación de una manera menos rotunda que el nivel de económico o el termino independiente.

También se puede observar una disminución en los coeficientes estimados, más notoria en el caso del nivel de riqueza. Originada por la combinación de los efectos que éste y la oferta de trabajo (directo e inverso respectivamente) tienen sobre los salarios y que se comprobaban en el *modelo 2*: si el nivel de salarios fuese variable explicativa, el no considerarla (como se hace en el *modelo 3*), haría que su efecto apareciese incluido en el de las otras variables. Así los coeficientes de los niveles de riqueza y educación estimados en el *modelo 1* son mayores por que no incluirían el supuesto efecto inverso del nivel de salarios.

Como conclusión se aceptaría para la variación de la oferta de trabajo una relación directa con la variación del nivel económico. Se aceptaría entonces que cuando la economía crece en términos reales la gente se animará a entrar en el mercado laboral. La explicación de este resultado vendría dada por el hecho de que si la economía está en una etapa expansiva aumentará el número de puestos de trabajo, con lo que disminuirá el coste de buscar y encontrarlos. Además, al incrementarse la demanda los empleadores ofrecerán mejores condiciones por lo que será mas fácil encontrar trabajos que superen el *salario de reserva* del individuo y lo animen a trabajar. De forma similar, en las épocas de recesión aumentan las dificultades de encontrar empleo y empeoran las condiciones en los empleos ofertados, esto hará aparecer el *efecto desánimo* y llevará a algunos individuos a retirarse del mercado laboral.

Por lo que respecta al nivel educativo, considerándolo como un índice del capital humano de la autonomía, se tendría una relación directa entre sus variaciones y las variaciones de la oferta laboral. El incremento en el nivel de estudios permitirá obtener mejores condiciones dentro del mercado laboral, lo que aumentará el interés de entrar o permanecer en él. Así se produciría una mayor integración de las mujeres en el mercado laboral, con lo que se compensaría el efecto, más instantáneo e inverso, que en la reducción de las tasas de actividad juvenil tiene el incremento en el número de estudiantes.

En conjunto, los *modelos 2 y 3* presentan un análisis causal donde el nivel educativo y la riqueza afectan a los salarios, tanto directamente como indirectamente a través de la oferta laboral. El *gráfico 1* presenta en esencia esta idea.

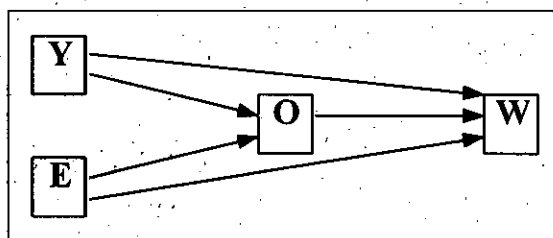


Gráfico 1

4.1.2 Estimación de los modelos con datos agregados.

Al hacer la estimación con los datos globales españoles la muestra será menor por lo que se esperarán unos resultados menos significativos. Otra diferencia que también se podría esperar a priori viene dada por el hecho de que, utilizando los datos obtenidos de las autonomías, se le está concediendo a todas la misma importancia, mientras que usando datos globales las autonomías con mayor población tendrán más peso en el resultado.

- Modelo 1:

$$O_t^* = -0.50250 + 0.010762 \cdot E_t + \mathbf{0.065074} \cdot Y_t - \mathbf{0.024028} \cdot W_t$$

(0.6284) (0.01237) (0.03102) (0.04237)

Los signos coinciden con los obtenidos en el modelo global, pero los coeficientes son aquí mayores; bastante mayores en el caso de los niveles de salarios y educación, y poco más en el caso del nivel de riqueza.

El coeficiente del nivel de educación no es significativo. El del nivel de riqueza sí, pero no de una forma tan rotunda como con los datos de las autonomías. El coeficiente del nivel de salarios es muy significativo, pero tiene signo negativo, igual que en el apartado anterior, por lo que se considerará esta variable como explicada por la oferta laboral y se estimarán los *modelos 2 y 3*.

- Modelo 2:

$$W_t^* = -6.3404 + 0.11553 \cdot E_t + \mathbf{1.6289} \cdot Y_t - \mathbf{12.777} \cdot O_t$$

(13.58) (0.2607) (0.7514) (3.050)

En estos resultados aparecen el nivel de riqueza y la oferta de trabajo como significativas en la explicación del modelo (la oferta de manera más rotunda que la riqueza), mientras que el nivel de educación no es significativo.

Los signos de los coeficientes coinciden con la estimación anterior, sin embargo sus valores son mucho mayores, llegando en el caso de la oferta de trabajo a ser 30 veces

mayor. La causa de esto está en la variabilidad de las variables, según se consideren datos autonómicos o estatales:

<u>Datos autonómicos</u>	<u>Datos estatales</u>	<u>Relación ^{Aut.}/_{Est.}</u>
$\sigma_W = 42.16$	$\sigma_W = 42.07$	1.002
$\sigma_E = 48.03$	$\sigma_E = 19.27$	2.492
$\sigma_Y = 10.53$	$\sigma_Y = 7.65$	1.376
$\sigma_O = 6.45$	$\sigma_O = 1.83$	3.525

Las diferencias entre la variabilidad autonómica y la estatal no son las mismas en todas las variables, oscilando entre la práctica igualdad del nivel de salarios, y la diferencia extrema de la oferta laboral, 3.5 veces más dispersos los datos autonómicos. Así, para variaciones iguales en los salarios, se necesitarán menores variaciones de la oferta de trabajo estatal (más estable) que de la oferta autonómica.

Otro aspecto que influye en las diferencias de los coeficientes estimados son las distintas dependencias que las variables autonómicas tienen con respecto a las correspondientes agregadas (*Cuadro 8*). Así, en los datos de los salarios (debido sobre todo a su construcción) R^2 oscila entre 0.888 (Canarias) y 0.987 (Valencia), con un valor promedio de 0.952, mientras que la oferta de trabajo oscila entre el 0.002 de La Rioja y el 0.202 de Castilla y León, con un promedio de 0.085. Debido a esto, una parte de la variación de los salarios autonómicos vendría determinada por la variación de los salarios estatales, lo que reduce la cantidad de variación explicada por las variables del modelo. Por el contrario, con los datos agregados esto no ocurre, por lo que en las variables afectadas los coeficientes tenderán a ser mayores con estos datos.

- Modelo 3:

$$O_t^* = -0.50527 + 0.011524 \cdot E_t + 0.037423 \cdot Y_t$$

(0.6984) (0.01425) (0.03273)

En este modelo ninguna de las dos variables resulta significativa, si bien el nivel de riqueza se acerca ($p=0.87$).

Las relaciones entre las variables en el modelo agregado son, en síntesis, las que se representan en el *gráfico 2*:

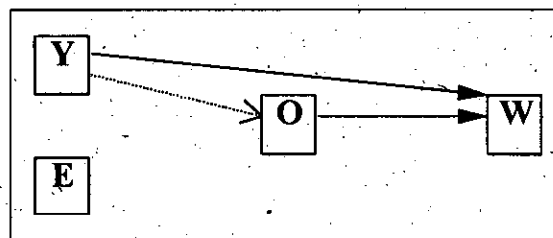


Gráfico 2

El nivel de educación parece la variable más afectada por las diferencias en la variabilidad. En cambio la variación del nivel de riqueza parece deberse más bien al efecto de las comunidades grandes en la estimación agregada (*cuadro 12*). Por lo que respecta a los signos de los coeficientes estimados, coinciden con los obtenidos usando datos autonómicos.

Resumiendo las observaciones hechas al comparar los resultados obtenidos en las dos estimaciones se tiene lo siguiente:

Los coeficientes son mucho más significativos al usar los datos autonómicos,

consecuencia de su mayor número.

Las estimaciones hechas con datos agregados se ven más influidas por las comunidades grandes.

Los cambios en la variabilidad de los datos, según sean agregados o autonómicos, difieren en las variables, afectándole a las estimaciones de los coeficientes.

Otro factor que influye en las estimaciones es la relación que tengan las variables autonómicas con las correspondientes estatales.

4.1.3 Estimación de los modelos en las autonomías.

Se hará ahora la estimación de los modelos para cada autonomía. Posteriormente se contrastará si alguna de ellas se diferencia significativamente del modelo global. En ese caso se eliminarán sus datos y se realizarán de nuevo las estimaciones de los

Valores del coeficiente de determinación entre los datos españoles y los datos autonómicos:							
Educación		Oferta		Salarios		P.I.B./hab.	
AND	0.677	AND	0.141	AND	0.952	AND	0.877
ARA	0.063	ARA	0.14	ARA	0.965	ARA	0.657
AST	0.437	AST	0.022	AST	0.957	AST	0.414
BAL	0.122	BAL	0.016	BAL	0.941	BAL	0.235
C_L	0.539	C_L	0.202	C_L	0.985	C_L	0.838
C_M	0.268	C_M	0.181	C_M	0.947	C_M	0.849
CAN	0.045	CAN	0.012	CAN	0.881	CAN	0.42
CAT	0.207	CAT	0.169	CAT	0.971	CAT	0.864
CNT	0.156	CNT	0.026	CNT	0.937	CNT	0.48
EXT	0.095	EXT	0.009	EXT	0.945	EXT	0.491
GAL	0.083	GAL	0.053	GAL	0.93	GAL	0.463
MAD	0.015	MAD	0.14	MAD	0.973	MAD	0.652
MUR	0.182	MUR	0.067	MUR	0.947	MUR	0.561
NAV	0.082	NAV	0.007	NAV	0.97	NAV	0.708
PV	0.052	PV	0.173	PV	0.983	PV	0.633
RIO	0.006	RIO	0.002	RIO	0.919	RIO	0.055
VAL	0.871	VAL	0.085	VAL	0.987	VAL	0.693
promedio: 0.229		promedio: 0.085		promedio: 0.952		promedio: 0.582	

Cuadro 8

modelos globales, comparando los resultados que se obtienen a partir de este nuevo grupo de datos, más homogéneos que los obtenidos a partir de los datos iniciales.

Se dispondrá para esta estimación de 43 datos en cada una de las comunidades. Los modelos estimados serán los siguientes⁷:

$$W = \beta_{0i} + \beta_{1i} O + \beta_{2i} E + \beta_{3i} Y \quad i=1, 2, \dots, 17 \quad (\text{modelo 2})$$

$$O = \theta_{0i} + \theta_{1i} E + \theta_{2i} Y \quad i=1, 2, \dots, 17 \quad (\text{modelo 3})$$

Siendo W el nivel de salarios,

O la oferta de trabajo,

E el nivel de educación,

Y el nivel económico de cada comunidad.

Se analizarán los resultados considerando un nivel de significación de 0.1 y dos colas. También se tendrán en cuenta los casos en los que $p > 0.9$ o $p < 0.1$, ya que, aunque no alcancen el nivel de significación, su cercanía puede ayudar a caracterizar los rasgos de las que si lo alcanzan.

⁷ Se estimó también el *modelo 1*: $O = \alpha_{0i} + \alpha_{1i} W + \alpha_{2i} E + \alpha_{3i} Y \quad i=1, 2, \dots, 17$

para comprobar si alguna comunidad se aparta del comportamiento global y en ella si puede ser considerado correcto. Por ese motivo los resultados de esta estimación aparecen en el *Apéndice IV.2*.

Los resultados para cada una de las autonomías son los siguientes:

Andalucía:

$$W_{t, \text{and}}^* = 0.32921 - 0.010508 \cdot E_{t, \text{and}} + 1.1758 \cdot Y_t - 1.9456 \cdot O_t$$

(7.008) (0.08373) (0.7226) (1.485)

$$O_{t, \text{and}}^* = -0.31560 + 0.012449 \cdot E_{t, \text{and}} - 0.080052 \cdot Y_{t, \text{and}}$$

(0.8021) (0.007890) (0.07622)

En esta autonomía ningún coeficiente llega a ser significativo. Están cerca de serlo las influencias de la oferta de trabajo y del nivel de riqueza sobre el nivel de salarios y la de nivel de educación sobre la oferta de trabajo.

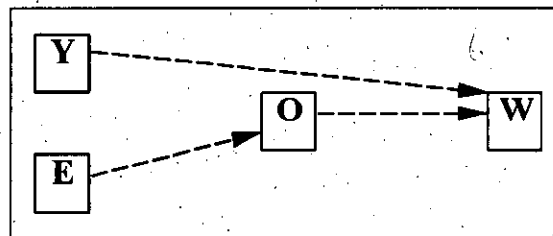


Gráfico 3

Aragón:

$$W_{t, \text{ara}}^* = -17.348 + 0.20673 \cdot E_{t, \text{ara}} + 3.0168 \cdot Y_{t, \text{ara}} - 3.4031 \cdot O_{t, \text{ara}}$$

(14.07) (0.2677) (0.7646) (1.295)

$$O_{t, \text{ara}}^* = -1.0884 + 0.0068716 \cdot E_{t, \text{ara}} + 0.16011 \cdot Y_{t, \text{ara}}$$

(1.810) (0.03421) (0.08089)

En esta comunidad la variación de los salarios aparece muy influida por la variación del nivel económico y por la variación de la tasa de actividad. En el *modelo 3* la variación del nivel económico también aparece como influencia significativa de la

variación de la oferta de trabajo. En conjunto, el nivel económico aparece en esta comunidad como una variable importante para el mercado de trabajo, tanto para los salarios como para la oferta. La situación opuesta es la de la variación del nivel de educación, que no resulta significativa en ninguno de los modelos.

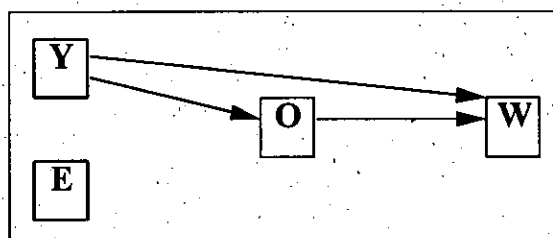


Gráfico 4

Asturias:

$$W_{t,ast}^* = -0.54268 - 0.045382 \cdot E_{t,ast} + 1.0360 \cdot Y_{t,ast} - 2.6686 \cdot O_{t,ast}$$

(7.677) (0.1375) (0.9292) (1.677)

$$O_{t,ast}^* = -1.1420 + 0.0038084 \cdot E_{t,ast} + 0.092634 \cdot Y_{t,ast}$$

(0.8422) (0.01367) (0.05858)

En Asturias ningún coeficiente estimado llega a ser significativo. Están cerca de serlo la influencia de la variación de la oferta de trabajo en la variación de salarios, y la variación del nivel de riqueza en la variación de la oferta laboral.

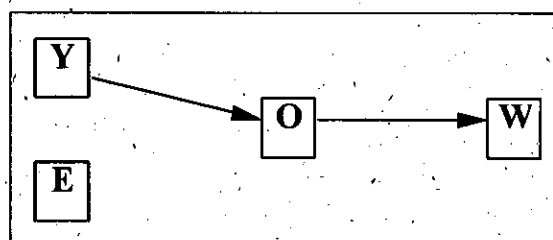


Gráfico 5

Baleares:

$$W_{t,bal}^* = 5.1153 - 0.045316 \cdot E_{t,bal} + 0.31548 \cdot Y_{t,bal} - 1.6689 \cdot O_{t,bal}$$

(8,879) (0,1395) (0,3107) (0,7705)

$$O_{t,bal}^* = 2.1768 - 0.042913 \cdot E_{t,bal} + 0.0054618 \cdot Y_{t,bal}$$

(1,944) (0,03146) (0,07625)

El único coeficiente significativo en esta comunidad es el de la variación de la oferta de trabajo en la variación de los salarios. Esta cerca de serlo la variación del nivel educativo en el *modelo 3*, además con un valor negativo, al contrario de las estimaciones globales realizadas. Signo negativo también lo tiene en ese modelo el coeficiente del nivel de riqueza, pero es prácticamente cero, de hecho en la estimación del *modelo 1* para esta comunidad (ver *Apéndice IV.2*) también sale muy cerca de cero pero con signo positivo, por lo que no se considerará este resultado. Sin embargo, el valor obtenido para el coeficiente del nivel de educación, negativo y casi significativo, puede hacer suponer que, en esta comunidad, el modelo de oferta de trabajo con estas variables sigue un comportamiento diferente al de las demás.

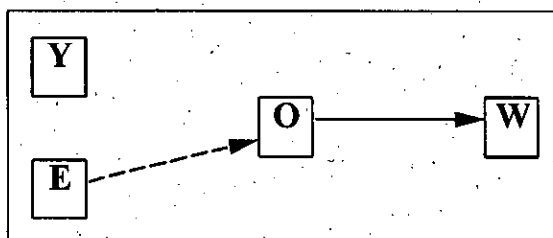


Gráfico 6

Canarias:

$$W_{t,can}^* = 1.5409 - 0.019596 \cdot E_{t,can} + 0.66617 \cdot Y_{t,can} - 0.90445 \cdot O_{t,can}$$

(7.313) (0.04318) (0.5667) (1.122)

$$O_{t,can}^* = 0.57583 - 0.0058631 \cdot E_{t,can} - 0.33269 \cdot Y_{t,can}$$

(1.328) (0.007246) (0.08213)

Al igual que en Baleares, en Canarias los coeficientes de las variables estimados en el *modelo 3* son negativos, al contrario de las estimaciones globales, sin embargo en este caso ninguno de los dos es significativo, como tampoco lo son los estimados en el *modelo 2*.

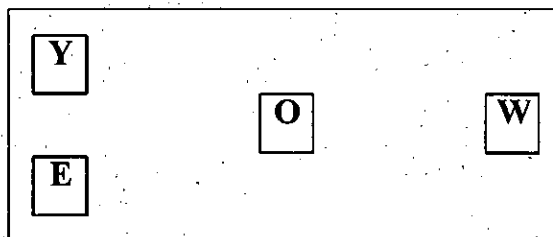


Gráfico 7

Cantabria:

$$W_{t,cnt}^* = -1.2897 + 0.035796 \cdot E_{t,cnt} + 0.94643 \cdot Y_{t,cnt} - 1.1758 \cdot O_{t,cnt}$$

(12.07) (0.2320) (0.6203) (1.011)

$$O_{t,cnt}^* = -3.4331 + 0.044231 \cdot E_{t,cnt} + 0.22986 \cdot Y_{t,cnt}$$

(1.879) (0.04067) (0.06477)

Aquí aparece el nivel económico como variable decisiva en los dos modelos, siendo muy significativa para la variación de la oferta de trabajo y casi significativa para la variación de los salarios. Las demás variables no tienen incidencia en estos modelos.

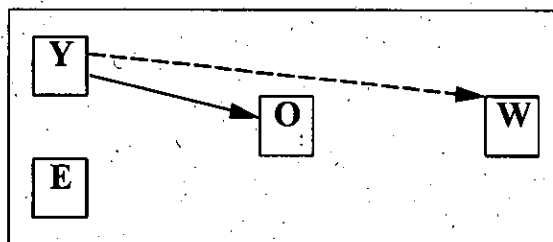


Gráfico 8

Castilla-La Mancha:

$$W_{t,cm}^* = -1.1153 + 0.057492 \cdot E_{t,cm} + 1.1602 \cdot Y_{t,cm} - 2.6079 \cdot O_{t,cm}$$

(6.514) (0.2163) (0.7311) (1.109)

$$O_{t,cm}^* = 0.030186 - 0.010214 \cdot E_{t,cm} + 0.11416 \cdot Y_{t,cm}$$

(1.135) (0.03403) (0.09365)

Ninguna variable sale significativa para la variación de la oferta de trabajo, sin embargo para la variación de los salarios sale el nivel de riqueza, y se acerca la variación de la oferta de trabajo.

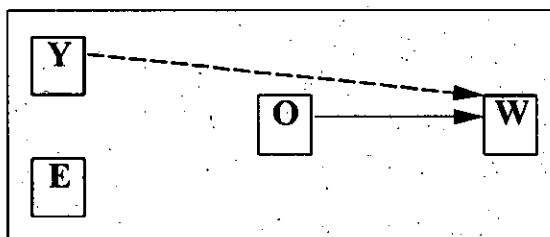


Gráfico 9

Castilla y León:

$$W_{t,cl}^* = 7.0268 - 0.43875 \cdot E_{t,cl} + 2.0232 \cdot Y_{t,cl} - 4.4430 \cdot O_{t,cl}$$

(11.36) (0.4483) (0.7100) (1.587)

$$O_{t,cl}^* = -1.4078 + 0.034542 \cdot E_{t,cl} + 0.063975 \cdot Y_{t,cl}$$

(1.239) (0.04854) (0.07726)

La oferta de trabajo y el nivel de riqueza son muy significativos en el *modelo 2*, pero ninguna variable lo es en el *modelo 3*.

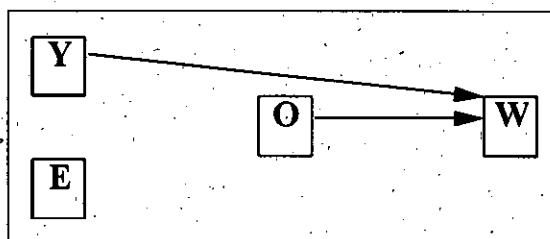


Gráfico 10

Cataluña:

$$W_{t,cat}^* = -2.4761 + 0.02081 \cdot E_{t,cat} + 1.1451 \cdot Y_{t,cat} - 1.353 \cdot O_{t,cat}$$

(10.89) (0.21) (0.5622) (2.8462)

$$O_{t,cat}^* = 0.36229 - 0.0067344 \cdot E_{t,cat} + 0.01564 \cdot Y_{t,cat}$$

(0.7905) (0.01493) (0.04087)

El único coeficiente significativo en esta comunidad es el de la variación de riqueza para la variación de los salarios.

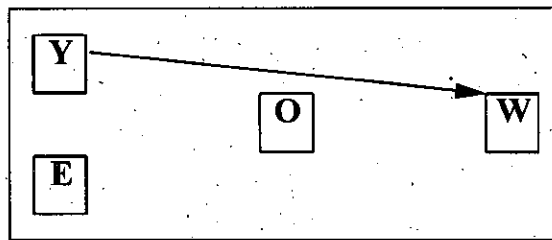


Gráfico 11

Comunidad Valenciana:

$$W_{t,val}^* = 0.75749 - 0.056697 \cdot E_{t,val} + 0.93493 \cdot Y_{t,val} + 0.41405 \cdot O_{t,val}$$

(9.874) (0.1613) (0.8180) (0.8406)

$$O_{t,val}^* = -0.33529 + 0.0094666 \cdot E_{t,val} + 0.0019785 \cdot Y_{t,val}$$

(0.9859) (0.01911) (0.09185)

En esta comunidad ningún coeficiente es significativo en ninguno de los dos modelos.

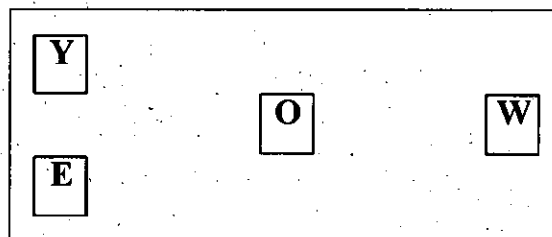


Gráfico 12

Extremadura:

$$W_{t,ext}^* = -2.8228 + 0.16037 \cdot E_{t,ext} + 0.61704 \cdot Y_{t,ext} - 1.2827 \cdot O_{t,ext}$$

(7.677) (0.1375) (0.9292) (1.677)

$$O_{t,ext}^* = -0.79140 + 0.039810 \cdot E_{t,ext} - 0.0084503 \cdot Y_{t,ext}$$

(0.5884) (0.01456) (0.08358)

Aquí la variable que parece importar más en el mercado de trabajo es la variación del nivel de educación, que resulta casi significativa para la variación de los salarios y significativa para la variación de la tasa de actividad.

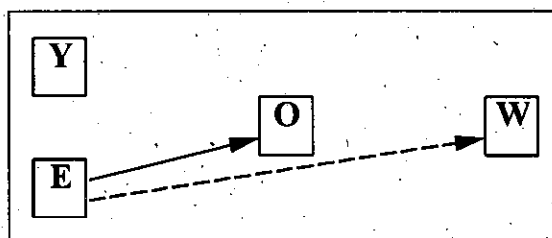


Gráfico 13

Galicia:

$$W_{t,gal}^* = -6.2260 + 0.053061 \cdot E_{t,gal} + 1.0201 \cdot Y_{t,gal} - 0.47416 \cdot O_{t,gal}$$

(8.414) (0.08347) (0.9899) (1.245)

$$O_{t,gal}^* = -1.7589 + 0.0057873 \cdot E_{t,gal} + 0.22975 \cdot Y_{t,gal}$$

(0.8422) (0.01367) (0.05858)

En Galicia ninguna variable parece llegar a afectar significativamente la variación de salarios, en cambio sí resulta significativa la variación del nivel de riqueza para la variación de la tasa de actividad.

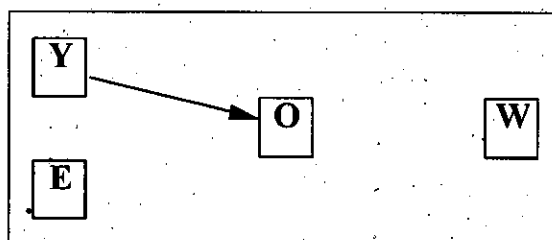


Gráfico 14

Madrid:

$$W_{t,mad}^* = -5.6831 + 0.090401 \cdot E_{t,mad} + 1.0400 \cdot Y_{t,mad} - 2.4296 \cdot O_{t,mad}$$

(13.31) (0.2597) (0.6075) (1.531)

$$O_{t,mad}^* = -0.073048 + 0.0014641 \cdot E_{t,mad} + 0.014591 \cdot Y_{t,mad}$$

(1.530) (0.06181) (0.02054)

Aquí resultan significativos para la variación de los salarios la variación del nivel de riqueza y la variación de la oferta de trabajo, sin embargo sobre la variación de la oferta de trabajo no parece actuar ninguna de las variables estudiadas, ya que los coeficientes son prácticamente cero.

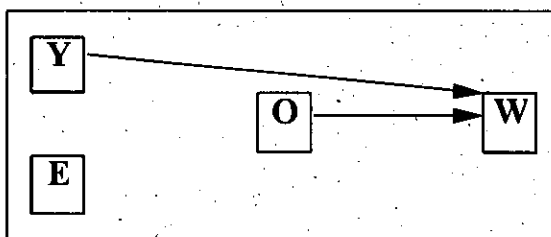


Gráfico 15

Murcia:

$$W_{t,mur}^* = -9.4955 + 0.10143 \cdot E_{t,mur} + 1.3910 \cdot Y_{t,mur} - 2.6375 \cdot O_{t,mur}$$

(5.607) (0.1073) (0.3560) (1.5473)

$$O_{t,mur}^* = -1.1966 + 0.022624 \cdot E_{t,mur} + 0.083180 \cdot Y_{t,mur}$$

(1.963) (0.02717) (0.08088)

Aquí, como en Madrid, también son significativos el nivel de riqueza y la oferta de trabajo para la variación de salarios, y ninguna lo es para la variación de la oferta de

trabajo. Sin embargo los coeficientes estimados para el *modelo 3* no están tan cerca de cero como lo estaban en Madrid.

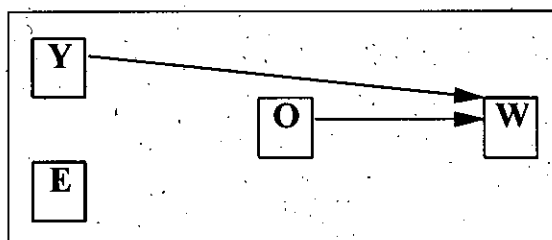


Gráfico 16

Navarra:

$$W_{t,nav}^* = -20.516 + 0.47275 \cdot E_{t,nav} + 1.1108 \cdot Y_{t,nav} - 0.17755 \cdot O_{t,nav}$$

(9.666) (0.2034) (0.4436) (1.056)

$$O_{t,nav}^* = 0.55219 - 0.030378 \cdot E_{t,nav} + 0.013029 \cdot Y_{t,nav}$$

(1.392) (0.02796) (0.06399)

La particularidad en esta autonomía es que la variación del nivel de educación parece actuar significativamente sobre la variación de los salarios, lo que no es habitual en las estimaciones globales ni en las otras autonomías (sólo en Extremadura se acerca). Además, en el *modelo 2*, también es significativo el nivel económico, lo que ya es más habitual, y en el *modelo 3* ningún coeficiente es significativo.

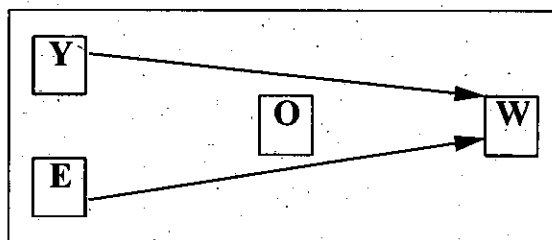


Gráfico 17

País Vasco:

$$W_{t,pv}^* = -3.9996 + 0.091920 \cdot E_{t,pv} + 0.83322 \cdot Y_{t,pv} - 4.0806 \cdot O_{t,pv}$$

(6.580) (0.1618) (0.8546) (1.423)

$$O_{t,pv}^* = -0.23171 - 0.0031105 \cdot E_{t,pv} + 0.079931 \cdot Y_{t,pv}$$

(0.6867) (0.01204) (0.03675)

En Euskadi resultan significativos la variación de la oferta de trabajo para los salarios y la variación del nivel económico para la oferta de trabajo.

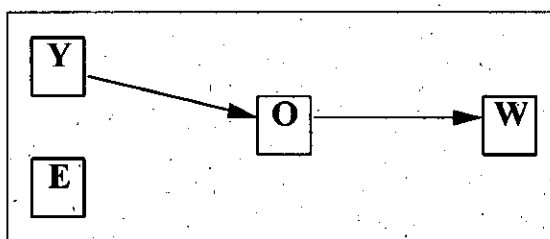


Gráfico 18

La Rioja:

$$W_{t,rio}^* = -0.54268 + 0.088739 \cdot E_{t,rio} + 0.52490 \cdot Y_{t,rio} + 0.99196 \cdot O_{t,rio}$$

(7.108) (0.1092) (0.4652) (0.5785)

$$O_{t,rio}^* = -2.1786 + 0.062063 \cdot E_{t,rio} + 0.0024828 \cdot Y_{t,rio}$$

(1.416) (0.02836) (0.07701)

En esta comunidad resulta significativa una relación directa entre variación de salarios y variación de oferta laboral, al revés que en las estimaciones globales. También resulta significativa la variación en el nivel educativo para la oferta de trabajo.

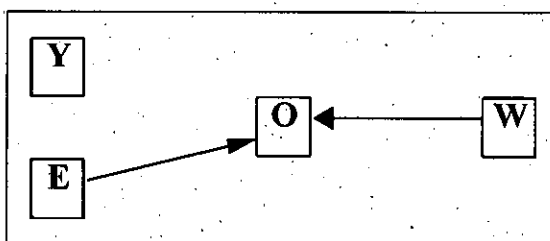


Gráfico 19

Analizando conjuntamente estos resultados y los obtenidos en la estimación con los datos agregados se podrían establecer excepciones y particularidades que ayuden a caracterizar los modelos.

En el *modelo 2* la significación de los coeficientes de las autonomías es bastante acorde con los resultados de España. Sin embargo en el *modelo 3*, donde ninguno de los coeficientes estimados con datos agregados resultó significativo, ya existe mayor disparidad.

La significación de estos coeficientes según las autonomías se puede ver en los *mapas 7 a 11*.

Las principales conclusiones que se pueden sacar del análisis de los resultados son las siguientes:

La variación del nivel de educación no resulta significativa en ninguno de los dos modelos ($p=0.67$ y $p=0.78$) ni en la mayoría de las autonomías (los valores de p están en su gran mayoría entre 0.2 y 0.8). Por lo tanto se podría considerar que excepto excepciones, esta variable no tiene influencia sobre las variaciones de salarios y de la oferta de trabajo. Las excepciones estarían en los casos siguientes:

Extremadura, donde resulta significativa para la oferta y se acerca en los salarios. Para la oferta también resulta significativa en La Rioja, con la cual coincide en ser una comunidad pequeña y con una fuerte dependencia de la agricultura (Raymond (1990)); está cerca de serlo Andalucía ($p=0.94$), cercana geográfica y económicamente (economía pobre y similares estructuras agrarias) a Extremadura.

Navarra, donde es significativa para los salarios y tiene un valor de $p=0.14$ y

coeficiente negativo para la oferta (contrario al habitual). Coeficiente negativo para la oferta también lo tienen los archipiélagos. En Canarias no está demasiado cerca de ser significativo ($p=0.21$), pero sí en Baleares ($p=0.09$).

La variación del nivel económico sí se puede considerar que influye en los salarios (el menor valor de p es 0.832 para el País Vasco), en cambio en la oferta de trabajo (*mapa 11*) sólo presenta significación (o casi) en los casos particulares de las comunidades de la Cornisa Cantábrica, donde la situación económica se deterioró fuertemente durante el periodo, y Aragón con un fuerte crecimiento (*cuadro 9*). Existen otras comunidades con fuertes crecimientos durante el periodo, como Madrid y los archipiélagos, sin embargo en ellas no resulta significativa esta relación. Lo que diferencia estas comunidades de Aragón fue el inicio de este crecimiento en un

Variación producida en las variables por autonomías entre los promedios de 1980 y 1990

Salarios	Educación	P.I.B./hab.	Oferta Laboral
CAN 13.73%	GAL 57.31%	CAN 35.89%	AND 4.11%
BAL 13.10%	VAL 43.55%	BAL 27.71%	MUR 2.95%
MAD 11.31%	AND 42.34%	EXT 24.41%	VAL 2.94%
ARA 10.07%	BAL 36.13%	MAD 23.31%	BAL 2.53%
CNT 9.08%	CAN 32.88%	ARA 22.30%	C_M 2.15%
EXT 7.64%	ESP 31.21%	C_M 18.85%	MAD 1.65%
C_L 5.32%	AST 28.75%	NAV 18.20%	EXT 1.53%
AST 5.01%	MUR 28.33%	MUR 16.08%	ESP 1.45%
NAV 4.23%	CAT 28.16%	ESP 13.54%	CAN 1.23%
ESP 4.22%	ARA 27.66%	CNT 12.82%	CAT 1.04%
CAT 3.34%	MAD 27.30%	CAT 12.34%	ARA 0.24%
PV 1.43%	CNT 25.87%	VAL 11.66%	RIO -1.88%
AND 0.92%	NAV 22.38%	RIO 10.75%	PV -2.61%
C_M 0.65%	C_M 21.83%	AND 8.30%	NAV -4.21%
RIO -0.95%	EXT 19.05%	C_L 8.27%	C_L -4.35%
VAL -1.74%	PV 17.10%	GAL 5.44%	CNT -5.33%
MUR -2.03%	RIO 14.56%	AST 4.55%	AST -6.22%
GAL -4.58%	C_L 14.18%	PV 1.51%	GAL -8.20%

Cuadro 9

acontecimiento singular, como fue la instalación de General Motors en Zaragoza, lo que crearía unas expectativas que animó a la gente a entrar en el mercado de trabajo. Influencia similar pero en sentido inverso tendría la crisis en la Cornisa Cantábrica, donde actuaría como un suceso impactante que desanimó la participación de gente en el mercado laboral.

En las demás comunidades no existe un efecto significativo de la variación del P.I.B. sobre la oferta de salarios. Sin embargo, trabajando con datos agregados (con menor variabilidad) esta relación presenta un valor $p=0.87$, lo suficientemente alto como para no descartar totalmente su existencia.

Respecto de la relación entre salarios y variación del nivel económico (*mapa 9*) también se puede observar que tiende a ser menos significativa en las comunidades en las que las variaciones fueron extremas durante el periodo (*Cuadro 9*). Esto se puede producir por la relación que tienen los datos de los salarios autonómicos y los agregados, así, la variación del salario en España podría actuar como un limitante para los salarios de las autonomías en las que la economía creció más fuerte, y un empuje en las economías con menor crecimiento.

La relación de la variación de la oferta de trabajo y los salarios es también bastante homogénea. Significativa en el caso español y en seis autonomías, y cerca de serlo en tres más (las autonomías son menos significativa por ser entidades de menor tamaño y por lo tanto su oferta de trabajo presenta mayores variaciones que en el caso global).

Una excepción a esta homogeneidad es La Rioja, que resulta ser significativa pero

con coeficiente negativo, al contrario del modelo global y del resto de las comunidades autónomas.

Las comunidades en las que esta variable tiene un valor de p muy bajo (entre 0.2 y 0.8), o sea, influencia nula en los salarios son: Galicia, Navarra, Cataluña, Valencia y Canarias. Estas comunidades están entre las que tienen mayores promedios de tasa de actividad durante el periodo (*Cuadro 10*) y, excepto Galicia, entre las de menor variabilidad. La única comunidad con estas características pero que tiene influencia significativa de esta variable es el País Vasco.

El comportamiento en estas comunidades lo explicaría su alta tasa de actividad y su poca variación, que impedirían, por un lado situaciones en las que la escasez de la oferta hiciese aumentar los salarios y, por otro, grandes excesos de oferta que superasen la resistencia de los salarios a bajar.

Estadísticos de la oferta de trabajo:

Oferta:	media	d.t.	1980	1990	diferencia	%
GAL	52,72%	1,49%	54,85%	50,36%	-4,50%	-8,20%
CAT	50,86%	0,72%	51,18%	51,71%	0,53%	1,04%
CAN	50,36%	0,87%	50,54%	51,16%	0,62%	1,23%
PV	50,25%	0,84%	51,98%	50,62%	-1,36%	-2,61%
VAL	50,16%	0,71%	49,72%	51,18%	1,46%	2,94%
BAL	49,39%	1,36%	50,57%	51,85%	1,28%	2,53%
NAV	49,13%	0,77%	50,84%	48,70%	-2,14%	-4,21%
RIO	49,05%	1,05%	50,38%	49,43%	-0,95%	-1,88%
ESP	48,81%	0,58%	49,11%	49,82%	0,71%	1,45%
CNT	48,22%	1,42%	50,72%	48,02%	-2,70%	-5,33%
MAD	48,11%	0,61%	48,03%	48,83%	0,79%	1,65%
AST	47,78%	1,12%	50,21%	47,09%	-3,12%	-6,22%
ARA	47,68%	0,69%	48,44%	48,56%	0,12%	0,24%
MUR	46,59%	0,86%	46,78%	48,16%	1,38%	2,95%
C_L	45,96%	0,91%	47,85%	45,77%	-2,08%	-4,35%
C_M	44,36%	0,74%	44,30%	45,25%	0,95%	2,15%
AND	44,34%	0,96%	43,91%	45,71%	1,80%	4,11%
EXT	44,27%	0,86%	44,48%	45,16%	0,68%	1,53%

Cuadro 10

La diferencia que presenta el País Vasco con respecto a las anteriores comunidades es la mayor fuerza y duración que la crisis tuvo en esa comunidad. En esa situación, un recurso que aplicaron las empresas para reducir el gasto salarial fueron las jubilaciones anticipadas, lo que actuaría de manera similar (sin serlo realmente) a un descenso de la oferta que permitiese a los salarios mantenerse a niveles más altos que lo que les correspondería por su situación económica.

En el caso de Galicia, la comunidad en la que la tasa de actividad tiene mayor desviación típica (*Cuadro 10*), la variabilidad procede de las reducciones sufridas por los sectores de la agricultura y la pesca tradicionales, donde los trabajadores son habitualmente autónomos. Así, esta reducción de la oferta laboral no se produciría en sectores que dependen del trabajo asalariado, por lo que los salarios no notarían su efecto.

Coeficientes estimados de las variables. En los modelos autonómicos, global (g) y agregado (e)

MODELO 2															
Gte				E				O				Y			
Aut	Coef. est.	t-ratio	P	Aut	Coef. est.	t-ratio	P	Aut	Coef. est.	t-ratio	P	Aut	Coef. est.	t-ratio	P
g	-3,1058	-2,108	0,018	CL	-0,43875	-0,9787	0,167	e	-12,777	-4,189	0	PV	0,83322	0,975	0,832
Nav	-20,516	-2,122	0,02	Can	-0,0196	-0,4539	0,326	Mur	-2,6375	-4,82	0	Ext	0,61704	0,9763	0,833
Mur	-9,4955	-1,694	0,049	Val	-0,0567	-0,3515	0,364	g	-0,4758	-9,708	0	Bal	0,31548	1,016	0,842
Ara	-17,348	-1,233	0,112	Ast	-0,04538	-0,3301	0,372	PV	-4,0806	-2,867	0,003	Gal	1,0201	1,03	0,845
Gal	-6,226	-0,7399	0,232	Bal	-0,04532	-0,3248	0,374	CL	-4,443	-2,799	0,004	Ast	1,036	1,115	0,864
PV	-3,9996	-0,6079	0,273	CM	-0,05749	-0,2658	0,396	Ara	-3,4031	-2,627	0,006	Rio	0,5249	1,128	0,867
Rio	-3,7942	-0,5338	0,298	And	-0,01051	-0,1255	0,45	CM	-2,6079	-2,351	0,012	Val	0,93493	1,143	0,87
e	-6,3404	-0,4668	0,322	Cat	0,02081	0,0991	0,539	Bal	-1,6689	-2,166	0,018	Can	0,66617	1,176	0,877
Ext	-2,8228	-0,4313	0,334	Cnt	0,035796	0,1543	0,561	Ast	-2,6686	-1,591	0,06	Cnt	0,94643	1,526	0,932
Mad	-5,6831	-0,4269	0,336	Mad	0,080401	0,3482	0,635	Mad	-2,4296	-1,587	0,06	CM	1,1602	1,587	0,94
Cat	-2,4761	-0,2274	0,411	e	0,11553	0,4432	0,67	And	-1,9456	-1,31	0,099	And	1,1758	1,627	0,944
CM	-1,1153	-0,1712	0,432	PV	0,09192	0,568	0,713	Cnt	-1,1758	-1,162	0,126	Mad	1,104	1,712	0,963
Cnt	-1,2897	-0,1088	0,458	Gal	0,053061	0,6356	0,736	Ext	-1,2827	-1,025	0,156	Cat	1,1451	2,037	0,976
And	0,32921	0,047	0,519	Ara	0,20673	0,7723	0,778	Can	-0,9045	-0,8058	0,213	e	1,6289	2,168	0,982
Ast	0,54268	0,07069	0,528	Rio	0,088739	0,8123	0,789	Cat	-1,353	-0,4754	0,319	Nav	1,1108	2,504	0,982
Val	0,75749	0,07672	0,53	Mur	0,10143	0,9452	0,825	Gal	-0,4742	-0,3807	0,353	CL	2,0232	2,85	0,997
Can	1,5409	0,2107	0,583	Ext	0,16037	1,324	0,903	Nav	-0,1776	-0,1681	0,434	g	0,44941	16,51	1
Bal	5,1153	0,5761	0,716	Nav	0,47275	2,325	0,987	Val	0,41405	0,4926	0,687	Mur	1,391	3,907	1
CL	7,0268	0,6187	0,73	g	0,018525	3,912	1	Rio	0,99196	1,715	0,953	Ara	3,0168	3,946	1

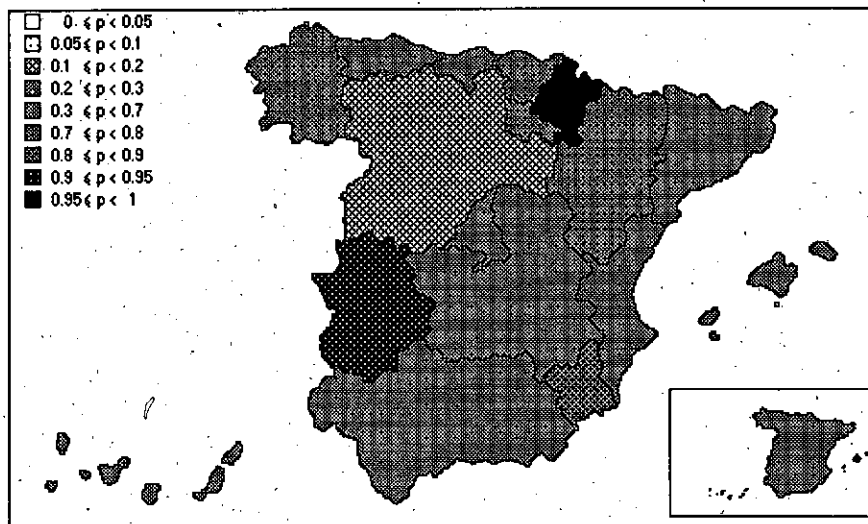
Cuadro 11

Coeficientes estimados de las variables. En los modelos autonómicos, global (g) y agregado (e)

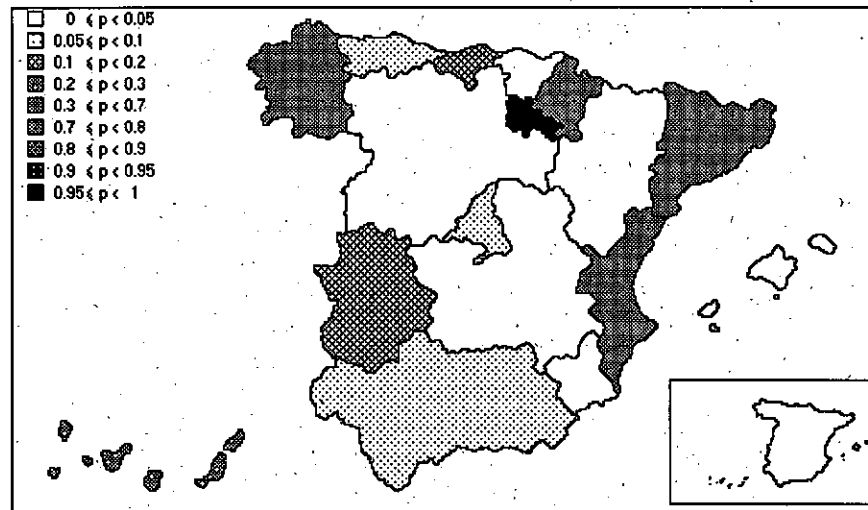
MODELO 3											
Cte				E				Y			
Aut	Coef.estm.	T-ratio	P	Aut	Coef.estm.	T-ratio	P	Aut	Coef.estm.	T-ratio	P
g	-0,79218	-3,818	0	Bal	-0,042913	-1,364	0,09	And	-0,080052	-1,05	0,2
Cnt	-3,4331	-1,827	0,04	Nav	-0,030378	-1,086	0,14	Can	-0,033269	-0,4051	0,3
Gal	-1,7589	-1,736	0,05	Can	-0,005863	-0,8092	0,21	Ext	-0,00845	-0,1011	0,5
Rio	-2,1786	-1,539	0,07	Cat	-0,006734	-0,4512	0,33	Bal	-0,005462	-0,07163	0,5
Ast	-1,142	-1,356	0,09	CM	-0,010214	-3,001	0,38	Val	0,0019785	0,02154	0,5
Ext	-0,7914	-1,345	0,09	PV	-0,003111	-0,2583	0,4	Rio	0,0024828	0,03224	0,5
CL	-1,4078	-1,137	0,13	Mad	0,0014641	0,07129	0,53	Nav	0,013029	0,2036	0,6
e	-0,50527	-0,7235	0,24	Ara	0,0068716	0,2009	0,58	Mad	0,014591	0,2361	0,6
Mur	-1,1966	-0,6097	0,27	Ast	0,0038084	0,2785	0,61	Cat	0,01564	0,3827	0,6
Ara	-1,0884	-0,6014	0,28	Gal	0,0057873	0,4724	0,68	CL	0,063975	0,8281	0,8
And	-0,3156	-0,3935	0,35	Val	0,0094666	0,4953	0,69	Mur	0,08318	1,028	0,8
Val	-0,33529	-0,3401	0,37	CL	0,034542	0,7117	0,76	e	0,037423	1,143	0,9
PV	-0,23171	-0,3374	0,37	e	0,011524	0,8085	0,79	CM	0,11416	1,219	0,9
Mad	-0,073048	-0,04774	0,48	Mur	0,022624	0,8328	0,8	Ast	0,092634	1,581	0,9
CM	0,030186	0,02659	0,51	Cnt	0,044231	1,088	0,86	Ara	0,16011	1,979	1
Nav	0,55219	0,3967	0,65	And	0,012449	1,578	0,94	PV	7,99E-02	2,175	1
Can	0,57583	0,4336	0,67	Rio	0,062063	2,188	0,98	Gal	0,22975	2,269	1
Cat	0,36229	0,4583	0,68	g	0,0059438	2,152	0,98	Cnt	0,22986	3,549	1
Bal	2,1768	1,12	0,87	Ext	0,03981	2,735	1	g	4,91E-02	3,621	1

Cuadro 12

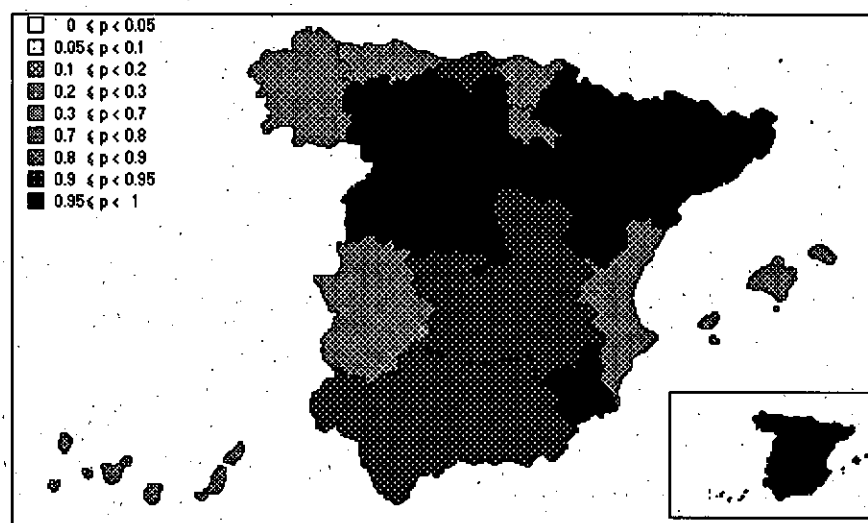
Significación de los coeficientes del MODELO 2:



Mapa 7: E

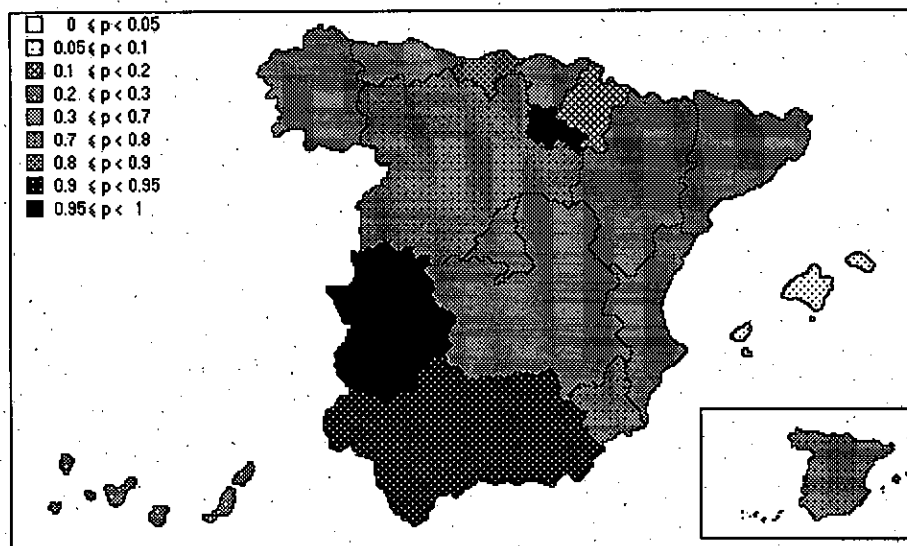


Mapa 8: O

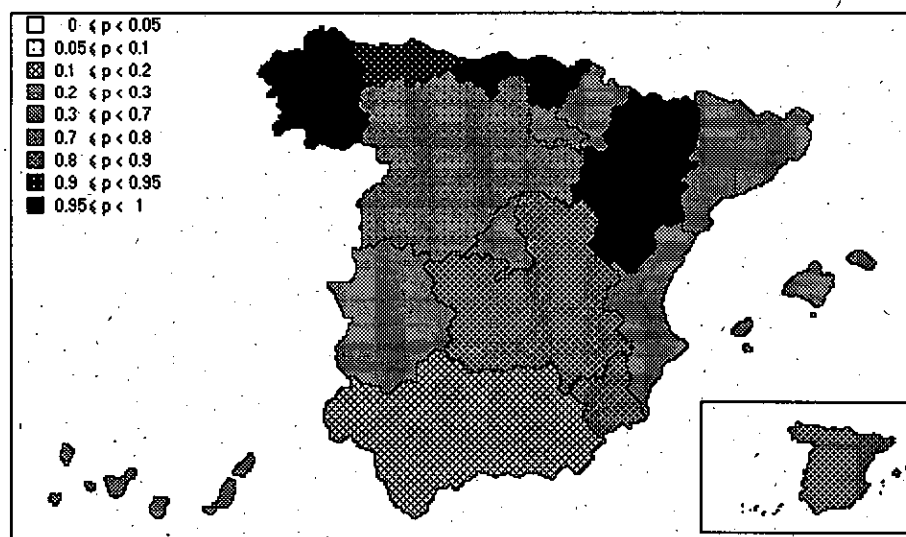


Mapa 9: Y

Significación de los coeficientes del MODELO 3:



Mapa 10: E



Mapa 11: Y

4.1.4 Contraste de diferencias en las comunidades.

Para contrastar si las diferencias entre los modelos autonómicos y el global son significativas se empleará el test de especificación de Hausman, que se describe en el *Apéndice IV.1*.

En el caso estudiado consideraremos las siguientes hipótesis de especificación:

H_0 : El modelo global es válido para todas las autonomías.

H_{1j} : El modelo de la autonomía j es diferente del global.

Bajo H_0 el estimador del modelo global será consistente, asintóticamente normal y asintóticamente insesgado y el estimador calculado con los datos de una autonomía será consistente y asintóticamente normal, pero no eficiente, ya que está calculado con menos datos que el del modelo global. Sin embargo, si se acepta H_{1j} el estimador de la autonomía seguirá siendo consistente, pero no así el global, que tenderá al modelo global de las demás autonomías, si existe, o a un promedio de todas ellas, en caso contrario.

Entonces para el *modelo 2* el contraste se distribuirá como una χ^2_4 , pues cuatro son los parámetros que se estiman en ese modelo. En el *modelo 3* en cambio será una χ^2_3 , ya que eliminando el nivel de salarios del modelo se tiene un parámetro menos para estimar.

Los resultados (que aparecen con detalle en el *Apartado IV.2*) muestran un *modelo 2* bastante heterogéneo entre las autonomías, ya que seis comunidades que se apartan significativamente del modelo estimado conjuntamente y tres más están cerca de hacerlo. El *modelo 3* es más homogéneo entre las distintas comunidades y el resultado global.

4.1.5 Estimación del modelo con comunidades homogéneas.

A partir de los contrastes de homogeneidad del modelo en las comunidades se pueden eliminar los datos de aquellas que sean significativamente más heterogéneas y volver a realizar la estimación global con el nuevo grupo de datos, en el intento de conseguir una mejor estimación del modelo global.

La heterogeneidad del *modelo 2* sugeriría la estimación de, por lo menos, dos modelos globales con dos grupos diferentes de comunidades, con lo cual los datos autonómicos no serían ya apropiados para estimar un modelo español para la variación de los salarios.

Los resultados del *modelo 3* son más homogéneos. Para un nivel de significación $\alpha=0.1$ resultarían dos comunidades heterogéneas: Cantabria y Extremadura. Una vez eliminados sus datos se obtendría la siguiente estimación para el modelo global:

- **Modelo 3:** (2ª estimación)

$$O_t^* = - \underset{(0.2260)}{0.672} + \underset{-(0.003101)}{0.0041988} \cdot E_t + \underset{(0.01511)}{0.039615} \cdot Y_t$$

Al eliminar los datos de Extremadura, una comunidad donde resultó significativo el coeficiente del nivel de educación, se redujo la significación de ese coeficiente a nivel global, pasando a tener un valor $p=0.912$.

Realizando nuevamente el test de Hausman para este modelo global y las comunidades con las que fue construido se obtienen que sólo una comunidad, Andalucía, es significativamente diferente.

Eliminados sus datos y reestimado el modelo con las comunidades restantes se obtiene la siguiente estimación:

- **Modelo 3:** (3ª estimación)

$$O_t^* = - \mathbf{0.63234} + 0.0019009 \cdot E_t + \mathbf{0.042979} \cdot Y_t$$

(0.2416) (0.003795) (0.01586)

Después de esta estimación el contraste no considera ninguna comunidad lo suficientemente distinta. Por lo que obtendríamos, una vez eliminados las comunidades más discordantes, un modelo global en el cual dejó de ser significativo el nivel de educación pero lo sigue siendo la variación del nivel económico. Este resultado es más coherente con el observado al estudiar las estimaciones autonómicas y de datos agregados, donde se concluía que el nivel de educación no parecía tener influencia sobre la oferta, mientras que la variación del nivel de riqueza parecía más factible que lo tuviese.

Respecto de los valores que toman los coeficientes estimados, son más cercanos al modelo global los obtenidos en la segunda estimación (con la excepción del término independiente que es más cercano en la tercera) esto se debe a la importancia que su tamaño concede a Andalucía en el cálculo de los valores agregados de España, por lo que, al proceder sus datos de un modelo significativamente diferente al de las demás comunidades, su eliminación influye alejando el resultado final del resultado agregado.

4.1.6 Conclusiones.

Teniendo en cuenta las diferentes estimaciones realizadas se puede concluir lo siguiente:

De las variables que se consideraron para explicar el comportamiento de los salarios, actúan sobre ellos las variaciones del nivel económico y las de la tasa de actividad.

Las variaciones del nivel de educación medidas por la proporción de estudiantes de secundaria no parecen tener ningún efecto sobre la variación de los salarios.

La relación observada entre nivel de salarios y nivel económico es directa: se esperan incrementos en los salarios cuando mejora la economía. Sin embargo, a nivel autonómico esta influencia parece amortiguada por la dependencia que los salarios autonómicos tienen de los salarios españoles.

La relación observada entre nivel de salarios y tasa de actividad es inversa: incrementos en la oferta de trabajo tienden a reducir los salarios. Observando las peculiaridades de los modelos autonómicos para situaciones en las que la tasa de actividad es alta y estable, las variaciones de la oferta de trabajo no se transmiten a los salarios. Tampoco se transmitió en el caso gallego al producirse principalmente la variación de la tasa de actividad en sectores no asalariados.

Para explicar la variación de la oferta de trabajo se descartó la variación de salarios, puesto que los resultados obtenidos en la primera estimación del modelo sugerían que esta influencia, de existir, tardaría un tiempo en manifestarse, por lo que no sería recogida en datos trimestrales, y la relación que se observa entre ambas variables se corresponde más bien a la influencia de la oferta en los salarios.

La variación del nivel económico sí actúa sobre la oferta a nivel global, con una relación directa, sin embargo las oscilaciones que la tasa sufre a nivel autonómico, mayores que a nivel agregado, ocultan su posible influencia, que no resulta significativa en la mayoría de las comunidades, manifestándose sólo cuando la situación económica fue afectada, positiva o negativamente, por algún acontecimiento excepcional y que afectase a la percepción que los posibles ofertantes de trabajo tienen sobre ella, animándolos o desanimándolos a participar en el mercado de trabajo. Este sería el caso de Aragón, con la instalación de la General Motors en Zaragoza (positivo), y de las comunidades de la Cornisa Cantábrica, donde jugó ese papel la fuerza con la que se vieron afectadas por la crisis económica.

Una vez estimados los modelos se realizó un contraste para estudiar posibles diferencias entre comunidades y el modelo global.

El contraste sobre el modelo de salarios aceptó seis comunidades como diferentes y tres más como cerca de serlo.

El contraste sobre el modelo de oferta de trabajo dio mayor homogeneidad entre las autonomías, excluyendo como diferentes sólo a Extremadura y Cantabria. Un segundo contraste, hecho sobre el modelo que se estimó eliminando las comunidades

anteriores dio a Andalucía como única comunidad diferente.

Haciendo finalmente la estimación global con datos de las comunidades homogéneas, se obtiene un resultado más coherente que el observado en los resultados agregado y autonómicos. Sin embargo la estimación que más se acercó a los valores obtenidos en la estimación agregada, fue la realizada eliminando sólo Extremadura y Cantabria, pero no Andalucía. Este hecho nos indica la importancia de las comunidades grandes a la hora de realizar la estimación de un modelo económico para España, debido a la cual, si una parte significativa del territorio español sigue un modelo diferente del mayoritario, su estimación agregada puede dar lugar a un modelo intermedio que no refleje la situación real.

Por otro lado, si estamos interesados en estimar un modelo global a partir de datos autonómicos, la eliminación de una comunidad grande por ser diferente va a apartar la estimación de lo que se obtendría con datos agregados. De la misma manera, la no eliminación de comunidades pequeñas diferentes también alteraría el resultado, puesto que se les estaría dando más importancia de la que en realidad les correspondería.

Otras características de las variables autonómicas que afectarán a la estimación global son su variabilidad y su dependencia de la situación española.

Otras características de las variables autonómicas que afectarán a la estimación global son su variabilidad y su dependencia de la situación española. Esto último provocará que los parámetros estimados en el modelo autonómico sean menores que en el agregado, ya que una parte de su variabilidad vendrá determinada por su situación a nivel español, que no aparece reflejada en el modelo. Esta situación es la que se presenta en las estimaciones del *modelo 2*.

4.2 ESTIMACIÓN SEMIPARAMÉTRICA DE LA OFERTA DE TRABAJO.

El último apartado de este capítulo está dedicado a la estimación de los modelos de la oferta de trabajo mediante suavización no paramétrica previa.

Como ya se comentó previamente, este método de estimación presenta dos decisiones importantes para obtener estimadores óptimos según la muestra observada, aunque sus propiedades teóricas son comunes:

- Selección del parámetro ventana.
- Selección del estimador no paramétrico previo de la regresión

Este apartado se va a dividir en tres partes. La primera estará dedicada a comentar el método de elección automática del parámetro ventana a partir de los valores de la muestra. A continuación se construirá el estimador, para lo cual se hace uso, como estimador no paramétrico previo, de un estimador Kernel multidimensional. Finalmente se aplica este estimador a los modelos considerados en el *Apartado 4.1*, haciendo uso de los datos de las autonomías y comparando esos resultados con los que se obtienen por Mínimos Cuadrados Ordinarios, ya que en las hipótesis consideradas para la construcción del nuevo estimador se exigía la homocedasticidad del modelo.

4.2.1 Elección automática del parámetro ventana.

Está poco estudiada la aplicación de métodos de búsqueda automática del parámetro ventana óptimo a la estimación mediante suavización no paramétrica previa.

Comentarios sobre el tema aparecen en González (1988) y en ellos se sugiere que

estos métodos deberán basarse en algún mecanismo de validación cruzada ("cross validation").

La idea básica de todo algoritmo de selección de la ventana es minimizar algún criterio de error entre la función estimada y la real. Algunos de esos criterios serían los siguientes:

Error cuadrático ponderado (ASE):

$$d_A(h) = \frac{\sum_{i=1}^n [\hat{\alpha}_n(X_i, h) - \alpha(X_i)]^2 \omega(X_i)}{n}$$

siendo $\alpha(X_i)$ la función de regresión teórica que se pretende estimar, $\hat{\alpha}_n(X_i, h)$ la función de regresión estimada, h el parámetro ventana y $\omega(X_i)$ una función de ponderación para evitar el efecto frontera.

Error cuadrático ponderado integrado (IASE):

$$d_I(h) = \int_S [\hat{\alpha}_n(x) - \alpha(x)]^2 \omega(x) f(x) dx$$

Error cuadrático ponderado medio (MASE):

$$\text{Condicionado: } d_C(h) = E \left[\frac{d_A(h)}{(X_1, X_2, \dots, X_n)} \right]$$

$$\text{Sin condicionar: } d_M(h) = E[d_A(h)]$$

entre estos criterios existe una equivalencia asintótica (Härdle (1990)), por lo cual, los óptimos que se obtienen con cada uno de ellos tienden a ser semejantes. A la hora de elegir uno de ellos, se tendrá en cuenta que nos proporcione un estimador piloto no paramétrico lo más ajustado posible a los datos (Härdle, Hall, Marron (1988)), por lo que se usaría el ASE como criterio.

Un estimador para $d_A(h)$ sería el error de predicción,

$$p(h) = \frac{\sum_{i=1}^n [Y_i - \hat{\alpha}_n(X_i, h)]^2}{n}$$

sin embargo va a ser un estimador sesgado. La razón intuitiva para este sesgo sería el uso de Y_j para predecirse a si mismo.

Existen varias maneras de construir un estimador insesgado de $d_A(h)$. La que se va a considerar aquí será la de validación cruzada, la cual predice los Y_j mediante estimadores de la regresión contruidos a partir de submuestras $\{(X_i, Y_i)\}_{i \neq j}$, con lo cual se evita el efecto de que Y_j se estime a si misma y el sesgo que eso produciría.

Por lo tanto, el estimador de $d_A(h)$ tomará la forma

$$CV(h) = \frac{\sum_{i=1}^n [Y_i - \hat{\alpha}_n^{(i)}(X_i, h)]^2}{n}$$

siendo $\hat{\alpha}_n^{(i)}(X_i, h)$ la función estimada a partir de la submuestra $\{(X_j, Y_j)\}_{j \neq i}$, mediante el método para el que se busca la ventana óptima.

Aplicada esta expresión a la estimación con suavización no paramétrica se tendría:

$$CV(h) = \frac{\sum_{i=1}^n [Y_i - A'(X_i) \hat{\Theta}^{(n)}(h)]^2}{n}$$

siendo $\hat{\Theta}^{(n)}(h)$ el parámetro estimado a partir de la submuestra $\{(X_j, Y_j)\}_{j \neq i}$.

Entonces, para unos datos dados $\{(X_i, Y_i)\}_{i=1}^n$ el valor de h que minimizase $CV(h)$ sería el valor que se tomaría como ventana óptima para la estimación de la regresión mediante suavización no paramétrica previa.

En Härdle (1990) aparece una descripción de métodos de búsqueda automática del parámetro ventana aplicados a la regresión no paramétrica. Una descripción más global de estos métodos aparece en Quintela (1992)

4.2.2 Diseño del Método de Estimación Semiparamétrica de la Oferta de Trabajo.

Para realizar la aplicación de este método a los datos se necesita escoger previamente una función δ con la que trabajar. La función δ escogida será un Kernel multidimensional, pues es la más usada entre las que cumplen las condiciones de las funciones δ :

$$\delta(x, X_k) = \prod_{i=1}^p \frac{1}{h_i} K\left(\frac{x^i - X_k^i}{h_i}\right)$$

Siendo $K(u)$ la siguiente función Kernel:

$$K(u) = \begin{cases} \frac{1}{2} & -1 < u < 1 \\ 0 & \text{en el resto} \end{cases}$$

Al ser indiferente el Kernel escogido para el resultado de esta estimación (Faraldo, González (1987)) se ha escogido este Kernel por su simplicidad para el cálculo.

Entonces tenemos que la función δ va a ser:

$$K(u) = \begin{cases} \frac{1}{2^p \prod_{i=1}^p h_i} & X_k^i - h_i < x^i < X_k^i + h_i \\ 0 & \text{en el resto} \end{cases}$$

Una vez elegida la función δ es necesario calcular las matrices C y D , calculando los diferentes elementos \bar{V} e \bar{T} ⁸.

Las matrices obtenidas son las siguientes:

$$C = \begin{pmatrix} B_C^1 & 0 & 0 & b^1 \\ 0 & B_C^2 & 0 & b^2 \\ 0 & 0 & B_C^M & b^M \\ (b^1)^t & (b^2)^t & (b^M)^t & B_C \end{pmatrix} \quad D = \begin{pmatrix} B_D^1 \\ B_D^2 \\ B_D^M \\ B_D \end{pmatrix}$$

⁸ Los cálculos están hechos en el Apéndice IV.3.

Siendo B_c^r , b^r , y B_c los siguientes bloques:

$$B_c^r = \begin{pmatrix} \frac{1}{M} & \frac{\overline{X_r^2}}{M} & \frac{\overline{X_r^q}}{M} \\ \frac{\overline{X_r^2}}{M} & \frac{(\overline{X_r^2})^2 + k_2}{M} & \frac{\overline{X_r^q X_r^2}}{M} \\ \frac{\overline{X_r^q}}{M} & \frac{\overline{X_r^q X_r^2}}{M} & \frac{(\overline{X_r^q})^2 + k_q}{M} \end{pmatrix} \quad b^r = \begin{pmatrix} \frac{\overline{X_r^{q+1}}}{M} & \frac{\overline{X_r^p}}{M} \\ \frac{\overline{X_r^{q+1} X_r^2}}{M} & \frac{\overline{X_r^p X_r^2}}{M} \\ \frac{\overline{X_r^{q+1} X_r^p}}{M} & \frac{\overline{X_r^p X_r^p}}{M} \end{pmatrix}$$

Siendo $\overline{X_r^i} = \frac{\sum_{t=1}^T X_{rt}^i}{T}$

$$\overline{(X_r^i)^2} = \frac{\sum_{t=1}^T (X_{rt}^i)^2}{T}$$

$$\overline{X_r^i X_r^j} = \frac{\sum_{t=1}^T X_{rt}^i X_{rt}^j}{T}$$

$$k_i = \frac{h_i^2}{3}$$

$$B_c = \begin{pmatrix} \frac{(\overline{X^{q+1}})^2 + k_{q+1}}{M} & \frac{\overline{X^{q+1} X^{q+2}}}{M} & \frac{\overline{X^{q+1} X^p}}{M} \\ \frac{\overline{X^{q+2} X^{q+1}}}{M} & \frac{(\overline{X^{q+2}})^2 + k_{q+2}}{M} & \frac{\overline{X^{q+2} X^p}}{M} \\ \frac{\overline{X^p X^{q+1}}}{M} & \frac{\overline{X^p X^{q+2}}}{M} & \frac{(\overline{X^p})^2 + k_p}{M} \end{pmatrix}$$

Siendo $\overline{(X^i)^2} = \frac{\sum_{t=1}^T \sum_{m=1}^M (X_{tm}^i)^2}{TM}$

$$\overline{X^i X^j} = \frac{\sum_{t=1}^T \sum_{m=1}^M X_{tm}^i X_{tm}^j}{TM}$$

$$k_i = \frac{h_i^2}{3}$$

B_D^r y B_D los siguientes bloques:

$$B_D^r = \begin{pmatrix} \frac{\overline{Y_r}}{M} \\ \frac{\overline{Y_r X_r^2}}{M} \\ \frac{\overline{Y_r X_r^q}}{M} \end{pmatrix} \quad B_D = \begin{pmatrix} \overline{Y X^{q+1}} \\ \overline{Y X^{q+2}} \\ \overline{Y X^p} \end{pmatrix}$$

y siendo $\overline{Y_r} = \frac{\sum_{t=1}^T Y_{rt}}{T}$

$$\overline{Y_r X_r^i} = \frac{\sum_{t=1}^T Y_{rt} X_{rt}^i}{T}$$

$$\overline{Y X^i} = \frac{\sum_{t=1}^T \sum_{m=1}^M Y_{tm} X_{tm}^i}{TM}$$

4.2.3 Resultados.

El modelo será estimado a partir de los mismos datos con los que se realizó la estimación en el *Apartado 4.1*: 731 datos obtenidos de las diferencias entre 44 trimestres en las 17 comunidades autónomas, y 43 datos obtenidos de las diferencias de 44 trimestres para los datos agregados españoles⁹.

Los resultados obtenidos para cada modelo empleando la totalidad de los datos fueron los siguientes:

- Modelo 1:

Estimación semiparamétrica

$$O_t^* = -0.65776 + 0.00638 \cdot E_t + 0.06263 \cdot Y_t - 0.01853 \cdot W_t$$

(0.2670) (0.00431) (0.01801) (0.0045)

Estimación M.C.O.

$$O_t^* = -0.65801 + 0.00639 \cdot E_t + 0.06264 \cdot Y_t - 0.01853 \cdot W_t$$

(0.2670) (0.00431) (0.01801) (0.00450)

Los dos métodos dan prácticamente los mismos resultados, tanto para los parámetros estimados como para sus varianzas.

⁹ Para facilitar los cálculos y evitar en lo posible errores de aproximación se hicieron los siguientes cambios de escala:

$$\begin{aligned} W' &= W; \\ E' &= 10^2 \cdot E; \\ Y' &= 10^3 \cdot Y; \\ O' &= 10^3 \cdot O; \end{aligned}$$

En estas estimaciones los signos de los coeficientes son los mismos que los estimados en el *Apartado 4.1*: relación directa del nivel de riqueza e inversa del nivel de salarios sobre la oferta laboral. Sin embargo, la influencia del nivel de educación no es significativa, un resultado que antes sólo se obtenía cuando se eliminaban los datos de comunidades que diferían significativamente del modelo global.

- Modelo 2:

Estimación semiparamétrica

$$W_t^* = - 2.59167 + 0.03167 \cdot E_t + \mathbf{0.90143} \cdot Y_t - \mathbf{1.23285} \cdot O_t$$

(2.1851) (0.03517) (0.1443) (0.2993)

Estimación M.C.O.

$$W_t^* = - 2.5945 + 0.03173 \cdot E_t + \mathbf{0.90146} \cdot Y_t - \mathbf{1.23298} \cdot O_t$$

(2.1851) (0.03517) (0.1443) (0.2993)

En este modelo vuelven a coincidir los resultados. Influye directamente el nivel de riqueza sobre los salarios e inversamente la oferta de trabajo.

- Modelo 3:

Estimación semiparamétrica

$$O_t^* = - \mathbf{0.62388} + 0.00593 \cdot E_t + \mathbf{0.04699} \cdot Y_t$$

(0.2698) (0.00435) (0.01779)

Estimación M.C.O.

$$O_t^* = - \mathbf{0.62419} + 0.00594 \cdot E_t + \mathbf{0.04701} \cdot Y_t$$

(0.2698) (0.00435) (0.01779)

Se repite lo comentado para los otros modelos: coincidencia de resultados entre los dos métodos y no influencia de la educación en la oferta de trabajo.

Como consecuencia de esta coincidencia de resultados podemos suponer que las perturbaciones son normales, ya que el método semiparamétrico es asintóticamente eficiente (Th 2.4.) y se emplean una elevada cantidad de datos en la estimación de estos modelos, mientras que el de M.C.O. es eficiente bajo normalidad. Sin embargo, esta coincidencia de resultados no se va a producir en la estimación de los modelos con datos agregados, ya que sólo se emplean 43 datos, por lo que el método no paramétrico se acerca menos a la eficiencia.

Los resultados obtenidos en las estimaciones con los datos agregados fueron los siguientes:

- Modelo 1:

Estimación semiparamétrica

$$O_t^* = -0.47698 + 0.01025 \cdot E_t + 0.06460 \cdot Y_t - 0.02398 \cdot W_t$$

(0.6439) (0.01285) (0.03306) (0.00586)

Estimación M.C.O.

$$O_t^* = -0.50250 + 0.01076 \cdot E_t + 0.06507 \cdot Y_t - 0.02403 \cdot W_t$$

(0.6356) (0.01268) (0.03263) (0.0578)

No resulta significativa la influencia del nivel de educación ni el término independiente en ninguno de los dos métodos. Sí resultan significativos el nivel de riqueza (influencia directa) y el nivel de salarios (influencia inversa). Los parámetros estimados semiparamétricamente son más cercanos a cero que los estimados por M.C.O. y su varianza es algo mayor.

- Modelo 2:

Estimación semiparamétrica

$$W_t^* = - 5.08674 + 0.08960 \cdot E_t + \mathbf{1.61937} \cdot Y_t - \mathbf{12.72677} \cdot O_t$$

(14.94) (0.2983) (0.7554) (3.1144)

Estimación M.C.O.

$$W_t^* = - 6.3404 + 0.11553 \cdot E_t + \mathbf{1.6289} \cdot Y_t - \mathbf{12.777} \cdot O_t$$

(14.74) (0.2945) (0.7455) (3.074)

Se obtiene influencia significativa y directa del nivel de riqueza sobre el nivel de salarios, y significativa e inversa de la oferta de trabajo. Otra vez los estimadores semiparamétricos son más cercanos a cero que los de M.C.O.

- Modelo 3:

Estimación semiparamétrica

$$O_t^* = - \mathbf{0.47613} + 0.01095 \cdot E_t + \mathbf{0.03675} \cdot Y_t$$

(0.7634) (0.01523) (0.03837)

Estimación M.C.O.

$$O_t^* = -0.50527 + 0.01152 \cdot E_t + 0.03742 \cdot Y_t$$

$$(0.7635) \quad (0.01523) \quad (0.03837)$$

Se repiten las características de los modelos anteriores: nivel de educación no significativo y nivel de riqueza significativo, con signo positivo. Además los estimadores semiparamétricos vuelven a ser más cercanos a cero y sus varianzas mayores.

Los resultados de la estimación de los modelos autonómicos están en el *Apéndice IV.4*. Las diferencias observadas entre los dos métodos son las mismas que se observan en los modelos españoles: parámetros estimados más cercanos a cero y varianzas mayores.

En resumen, los resultados que se obtienen indican un sesgo en la estimación semiparamétrica que acerca sus valores a cero. Este sesgo es producido por la forma que el estimador toma al usar un estimador Kernel como estimador no paramétrico previo, y en principio no se podría generalizar a otros casos¹⁰.

Las diferencias son, sin embargo, muy pequeñas, no superando el 2% para los parámetros significativos, y el 5% para los que están cerca de serlo. Además, en el

¹⁰ Si consideramos el estimador semiparamétrico como:

$$\hat{\beta}^* = (X'X+D)^{-1}X'Y = (X'X+D)^{-1}X'X\beta + (X'X+D)^{-1}X'E$$

siendo D una matriz diagonal formada con los parámetros ventana, su esperanza será :

$$E[\hat{\beta}^*] = (X'X+D)^{-1}X'X\beta = [I-(X'X+D)^{-1}D]\beta$$

caso de la estimación global (con más datos) los resultados son prácticamente coincidentes, debido a la eficiencia asintótica del estimador semiparamétrico.

En variables en las que el parámetro estimado por Mínimos Cuadrados es muy pequeño, también se observa que el parámetro ventana óptimo es muy grande, de forma que anula prácticamente la variación de la variable y estima semiparamétricamente un coeficiente cero.

5. CONCLUSIONES.

Como resumen del trabajo se resaltan las siguientes conclusiones:

- Se construye un estimador semiparamétrico para la regresión en datos de panel mediante suavización no paramétrica previa, (*apdo. 2.3.1*), extendiendo el modelo propuesto para la regresión múltiple por Cristóbal, Faraldo, González (1987). Esto permitirá analizar conjuntos de datos de panel, como es el caso de modelos económicos con datos de comunidades autónomas, sin hacer suposiciones sobre la ley de distribución de los errores. Además se demuestra que el estimador verifica la consistencia casi segura cuando se emplean los estimadores δ en la suavización no paramétrica previa. (*Apartado 2.3.2 y apéndice II.2*) y a continuación se obtiene su distribución asintótica (*Apartado 2.3.2 y apéndice II.2*).
- La metodología anterior se aplica a un modelo para la oferta de trabajo en las comunidades autónomas, considerando su relación con los niveles de salarios, educación y actividad económica de cada comunidad. (*Apdo. 3.3.1*)
- En las comunidades de salarios más bajos parece existir una relación entre la variación del nivel de salarios en el periodo y la variación del nivel educativo (porcentaje de alumnos de enseñanza secundaria) (*Apdo. 3.3.1*), lo que concuerda con la importancia del coste de oportunidad a la hora de decidir la prolongación de los estudios.

- Las series de las variables estudiadas (nivel de salarios, nivel de educación, P.I.B. por habitante y tasa de actividad), tanto a nivel desagregado en cada comunidad autónoma como en el global de España, se comportan como series de tendencia estocástica (*Apéndice III.4*), por tanto su comportamiento es impredecible a largo plazo.
- En particular, por ser el principal objeto de estudio se analizó la oferta de trabajo, medida mediante las series de diferencias trimestrales de las tasas de actividad autonómicas y españolas (*apdo. 3.5*). Mayoritariamente se comprobó que las series de las diferencias de la oferta de trabajo se comportan como ruido blanco.
- No obstante, en algunos casos se presentaban las particularidades siguientes:
 - * Estacionalidad, en seis comunidades y en la serie española, relacionada con el turismo (en Baleares, Canarias y posiblemente también en Andalucía) y con la estacionalidad de la agricultura de las zonas secas (Aragón, Castilla-La Mancha y Castilla y León).
 - * Un modelo MA(1) para las diferencias en cuatro comunidades (Navarra, Valencia, Castilla-La Mancha y Cantabria).
 - * Durante el análisis previo de las desviaciones de la serie también se encontró que la mayoría de ellas estaban relacionadas con dos hechos: el cambio metodológico de la E.P.A., comenzado a aplicar en el segundo trimestre de 1987; y el relanzamiento del empleo que se produce en algunas comunidades al salir de la crisis.

- Como consecuencia del análisis previo se consideraron tres posibles modelos:
 - * Un modelo global, estimado con los datos procedentes de las diferencias trimestrales de las variables autonómicas (*apdo. 4.1.1*).
 - * Un modelo español agregado, estimado con los datos procedentes de las diferencias trimestrales de las variables españolas (*apdo. 4.1.2*).
 - * Diecisiete modelos autonómicos, estimados con los datos procedentes de las diferencias trimestrales de las variables de la autonomía correspondiente (*apdo. 4.1.3*).
 - Los resultados obtenidos en la estimación del modelo global (*apartado 4.1.1*) llevan a suponer que el modelo más coherente con la teoría económica es aquel en que la variación del nivel de salarios depende de la oferta de trabajo y no la inversa. En consecuencia se realizó la estimación de un modelo que intenta explicar la variación de salarios mediante las variaciones del nivel de educación, del nivel económico y de la oferta de trabajo. Los resultados de esta estimación permiten aceptar influencias significativas para las tres variables: directa en los casos del nivel educativo y del económico e inversa en el caso de la oferta de trabajo.
- Para analizar las variables que afectan a la oferta laboral se estimó un modelo sin el nivel de salarios como variable explicativa. Los resultados permiten aceptar como significativas influencias directas del nivel de educación y del nivel económico sobre la oferta de trabajo.

- Los coeficientes estimados en el modelo agregado (*Apartado 4.1.2*), al emplear para su cálculo un menor número de datos, son menos significativos que en el modelo global.

Aquí los resultados también permiten suponer que es la oferta laboral la que actúa sobre la variación de los salarios, por lo que se estima primero un modelo para la variación de salarios, resultando significativos los coeficientes de la variación económica (relación directa) y de la oferta laboral (relación inversa). En cambio en el modelo agregado de la oferta de trabajo no resulta significativo ninguno de los coeficientes.

Comparando los resultados con los obtenidos en la estimación del modelo global, se aprecia el efecto de las distintas variabilidades de los datos según sean agregados o autonómicos, y que además cambian en distinta proporción para cada variable. También se aprecia el efecto que la relación de las variables autonómicas pueden tener con las correspondientes agregadas.

- Para los modelos autonómicos, a la vista de las estimaciones anteriores, se acepta a priori que la oferta de trabajo influye sobre la variación de los salarios, estimándose en cada autonomía un modelo para la variación de salarios y otro para la oferta de trabajo con sólo dos variables explicativas.

El menor número de datos empleado hace que se obtengan coeficientes menos significativos que en el caso global, además, también tienden a ser menos significativos que en el caso agregado, si bien aquí por causa de la mayor variabilidad de los datos autonómicos.

En la estimación del modelo autonómico de la variación de los salarios se obtienen altos niveles de significación para los coeficientes del nivel económico, que toman además valores positivos. De igual modo se observa que la dependencia que los salarios autonómicos tienen respecto de los salarios agregados actúa como suavizante frente a fuertes subidas o caídas que pudiese propiciar la situación económica de la autonomía.

Los coeficientes de la oferta de trabajo son menos homogéneos que los anteriores, resultando positivos en general, excepto en Valencia (donde no es significativo) y La Rioja. Este coeficiente es significativo en seis comunidades (País Vasco, Castilla-La Mancha, Castilla y León, Aragón, Murcia y Baleares) y cercana a serlo en tres más (Asturias, Madrid y Andalucía). Se observa en los resultados de este coeficiente que la variación de los salarios no se ve influida por la variación de la oferta de trabajo en comunidades con oferta laboral alta y estable y en Galicia, donde la crisis redujo puestos de trabajo sobre todo en sectores no asalariados.

Los coeficientes del nivel de educación no resultan significativos para este modelo, siendo Navarra la única excepción, mientras que Extremadura está cerca de serlo.

- Los modelos autonómicos de oferta laboral tienen pocos coeficientes significativos. En el nivel de educación sólo lo son Extremadura y La Rioja, acercándose Andalucía, mientras que para el nivel económico lo son los coeficientes de las autonomías de la Cornisa Cantábrica y de Aragón, comunidades en las que la situación económica vino determinada por sucesos

ocurridos en los primeros años del periodo (la crisis en la Cornisa Cantábrica y la instalación de la General Motors en Aragón), lo que pareció afectar a la percepción que los posibles ofertantes de trabajo tenían del mercado laboral.

- Los modelos estimados no son homogéneos en todas las autonomías, empleándose el test de Hausman para comprobar si las diferencias existentes llegan a ser significativas (*Apartado 4.1.4*), resultando que en el modelo de la variación de salarios seis comunidades son significativamente diferentes ($\alpha=0.1$), mientras que tres casi lo son ($0.1 < p < 0.2$).
- El modelo de la oferta de trabajo es más homogéneo, resultando significativamente diferentes dos comunidades, Extremadura y Cantabria, en un primer contraste y, reestimando el modelo global sin los datos procedentes de estas dos comunidades, también resulta diferente Andalucía.
- La estimación definitiva para el modelo global de oferta de trabajo se realiza con los datos procedentes de las comunidades homogéneas (*Apartado 4.1.4*). Aquí resulta significativa la influencia del nivel económico sobre la oferta laboral, pero no la variación del nivel educativo. Este resultado está más acorde con los obtenidos en los modelos autonómicos.
- El empleo de datos autonómicos para la estimación de un modelo económico proporciona una variedad de escenarios en los que estudiar el comportamiento del modelo. Otra ventaja que presenta trabajar con este tipo de datos es que podemos analizar si un modelo global es válido para todo el estado o tiene diferentes

comportamientos en partes importantes del territorio que tergiversarían los resultados obtenidos a partir de los datos agregados.

- Estimando un modelo global con datos autonómicos se obtiene mayor significación para los coeficientes estimados, ya que se emplea un mayor número de datos. Sin embargo, a la hora de sacar conclusiones de estos resultados se debe tener en cuenta el efecto de las mayores variabilidades de las variables autonómicas y de la dependencia que puedan tener respecto de las variables agregadas. Otra cosa que se debe tener en cuenta es que usando datos agregados se pondera más las comunidades con mayor población, mientras que usando todos los datos autonómicos todas tendrían la misma ponderación.
- Para estimar los modelos anteriores se diseñó una aplicación del estimador semiparamétrico, tomando como estimador δ previo un kernel multidimensional (*Apartado 4.2 y Apéndice IV.3*). Los resultados obtenidos en dicha estimación son semejantes a los de mínimos cuadrados, pero más cercanos a cero (a causa del sesgo). Esta semejanza aumenta en los modelos globales donde el número de datos empleados es mayor, lo que concuerda con la eficiencia asintótica de este estimador. Estos resultados permiten suponer además que no se puede rechazar el comportamiento normal de las perturbaciones.

APÉNDICE II: Demostración de las propiedades del estimador semiparamétrico.

APÉNDICE II.1

En este apéndice se citan algunos resultados que se emplearán en la demostración de los resultados de esta tesis.

Th. ("Desigualdad de Bennet") BENNET (1962)

Si $\{R_1, R_2, \dots, R_n\}$ es una muestra aleatoria simple de R , con R verificando:

i) $a \leq R \leq b$ con $a, b \in \mathbb{R}$

ii) $E(R)=0$ y $E(R^2)=\sigma^2$, entonces si $S_n = \sum_{i=1}^n R_i$ se verifica:

$$P\left\{\left|\frac{S_n}{n}\right| \geq \varepsilon\right\} \leq 2 \cdot e^{\left\{\frac{-n\varepsilon^2}{2[\sigma^2 + (b-a)\varepsilon]}\right\}}$$

Este resultado también es cierto si la muestra es de variables aleatorias

independientes con varianzas $\{\sigma_i^2\}$ con $i=1, \dots, n$, sin más que sustituir σ^2 por $\frac{\sum_{i=1}^n \sigma_i^2}{n}$.

Th: ("Criterio de Liapunov generalizado") ARAUJO-GINÉ (1980)

Si $Z_m = (Z_m^1, \dots, Z_m^m)$ es una sucesión de vectores aleatorios con $i = 1, \dots, k$ verificando:

$$i) \forall r, j=1, \dots, m; \sum_{i=1}^{k_n} E(Z_m^r Z_m^j) \rightarrow a_{rj}, \text{ con } A = (a_{rj}) \text{ matriz no nula.}$$

$$ii) \sum_{i=1}^{k_n} \int_{\|Z_m\| > \varepsilon} \|Z_m\|^2 dP \rightarrow 0 \quad \forall \varepsilon > 0,$$

entonces:

$$\sum_{i=1}^{k_n} Z_m dP \rightarrow N(0, A)$$

Th: ("Método δ ") BISHOP (1975)

Si $\hat{\theta}_n$ es un estimador vectorial de θ tal que verifica:

$$r_n(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, \Sigma(\theta)) \quad \text{con } r_n \rightarrow \infty$$

y g es una función definida en un abierto que contenga a θ y diferenciable continuamente en tal abierto, entonces:

$$r_n(g(\hat{\theta}_n) - g(\theta)) \xrightarrow{d} N\left(0, \left(\frac{\partial g}{\partial \theta}\right) \Sigma(\theta) \left(\frac{\partial g}{\partial \theta}\right)'\right)$$

APÉNDICE II.2

Este apéndice incluye las demostraciones de los resultados introducidos en este trabajo.

Demostración de la proposición (2.2).

Si hacemos el cambio $t=x-u$, tendremos $x=t+u$

$$dx=dt$$

Sustituyendo en la integral nos quedará:

$$\int g(x) \delta_s(u, x) dx = \int g(t+u) \delta_s(u, t+u) dt.$$

Como g es lineal:

$$\begin{aligned} \int g(x) \delta_s(u, x) dx &= \int [g(t)+g(u)] \delta_s(u, t+u) dt = \int g(t) \delta_s(u, t+u) dt + \int g(u) \delta_s(u, t+u) dt = \\ &= \Delta_1 + \Delta_2 \end{aligned}$$

Debido a que $g(u)$ no depende de t , tendremos que

$$\Delta_1 = g(u) \int \delta_s(u, t+u) dt$$

Además aplicando la propiedad (iii) de (2.1):

$$\Delta_2 = g(u)$$

$$\begin{aligned} \Delta_1 &= \int g(t) \delta_s(u, t+u) dt = \int g(x-u) \delta_s(u, x) dx = \\ &= \int_{\{\|x-u\| \leq c\varepsilon(n)\}} g(x-u) \delta_s(u, x) dx + \int_{\{\|x-u\| > c\varepsilon(n)\}} g(x-u) \delta_s(u, x) dx \end{aligned}$$

Por (iii), en la segunda integral $\delta_s(u, x) = 0$, por lo tanto esa integral también es igual a cero y

$$\Delta_1 = \int_{\{\|x-u\| \leq c\varepsilon(n)\}} g(x-u) \delta_s(u, x) dx$$

Teniendo en cuenta que $\|g(x-u)\| \leq \|x-u\| \leq c\varepsilon(n)$:

$$\|\Delta_1\| \leq \int_{\{\|x-u\| > c\varepsilon(n)\}} \|g(x-u)\| \delta_s(u, x) dx \leq \int_{\{\|x-u\| > c\varepsilon(n)\}} \|x-u\| \delta_s(u, x) dx$$

entonces:

$$\|\Delta_1\| \leq c\varepsilon(n) \int_{\{\|x-u\| > c\varepsilon(n)\}} \delta_s(u, x) dx \leq c\varepsilon(n) \rightarrow 0 \text{ cuando } n \rightarrow \infty$$

Por lo tanto $\Delta_1 \rightarrow 0$, y $\int g(x) \delta_s(u, x) dx = \Delta_1 + \Delta_2 \rightarrow g(u)$ cuando $n \rightarrow \infty$

c.q.d.

Demostración del lema (2.3).

(apartado 1):

Elemento \bar{T}_i^s :

Consideremos la descomposición

$$\bar{T}_i^s - T_{0,i}^s = \bar{T}_i^s - E[\bar{T}_i^s] + E[\bar{T}_i^s] - T_{0,i}^s = \Delta_1 + \Delta_2$$

Donde Δ_1 es la parte aleatoria y Δ_2 la parte determinística.

$$\begin{aligned} \Delta_2 &= \left\{ \frac{1}{TM} \sum_{i=1}^T \sum_{m=1}^M T_{i,m}^s \right\} - T_{0,i}^s = \\ &= E \left\{ \frac{1}{TM} \sum_{i=1}^T \sum_{m=1}^M \{ x^i I^s(x) Y_m \delta(x, X_m) dx \} \right\} - T_{0,i}^s = E \left\{ \int x^i I^s(x) Y \delta(x, X) dx \right\} - T_{0,i}^s \end{aligned}$$

Entonces:

$$E \left\{ \int x^i I^s(x) Y \delta(x, X) dx \right\} = E \left\{ E \left[\left(\int x^i I^s(x) Y \delta(x, X) dx \right) / X \right] \right\} =$$

$$E \left\{ \left(\int x^i I^s(x) \delta(x, X) dx \right) E[Y/X] \right\} = E \left\{ \left(\int x^i I^s(x) \delta(x, X) dx \right) \alpha(X) \right\}$$

Teniendo en cuenta que $\alpha(X) = \sum_{h=1}^q \sum_{i=1}^M \theta_h I^i(X) X^h + \sum_{q=h+1}^p \theta_h X^h$

$$\begin{aligned} &E \left\{ \left(\int x^i I^s(x) \delta(x, X) dx \right) \alpha(X) \right\} = \\ &= E \left\{ \left(\int x^i I^s(x) \delta(x, X) dx \right) \left(\sum_{h=1}^q \sum_{i=1}^M \theta_h I^i(X) X^h + \sum_{q=h+1}^p \theta_h X^h \right) \right\} = \end{aligned}$$

$$= \int \left(\int x' I^s(x) \delta(x, u) dx \right) \left(\sum_{h=1}^q \sum_{t=1}^M \theta_h I'(u) u^h + \sum_{q=h+1}^p \theta_h u^h \right) f(u) du$$

Usando la **proposición (2.2)** se deduce, por el *teorema de la convergencia dominada*, que:

$$\begin{aligned} E \left\{ \left(\int x' I^s(x) \delta(x, X) dx \right) \alpha(X) \right\} &\rightarrow \int u' I^s(u) \left(\sum_{h=1}^q \sum_{t=1}^M \theta_h I'(u) u^h + \sum_{q=h+1}^p \theta_h u^h \right) f(u) du = \\ &= \sum_{h=1}^q \sum_{t=1}^M \theta_h \left(\int u' I^s(u) I'(u) u^h f(u) du \right) + \sum_{q=h+1}^p \theta_h \left(\int u' I^s(u) u^h f(u) du \right) = \\ &= \sum_{h=1}^q \sum_{t=1}^M \theta_h E \left[x' I^s(x) I'(x) x^h \right] + \sum_{q=h+1}^p \theta_h E \left[x' I^s(x) x^h \right] = \\ &= \sum_{h=1}^q \theta_h E \left[x' I^s(x) x^h \right] + \sum_{q=h+1}^p \theta_h E \left[x' I^s(x) x^h \right] = \bar{T}_i^s \end{aligned}$$

Con esto se ve que $E \left\{ \left(\int x' I^s(x) Y \delta(x, X) dx \right) \right\} \rightarrow T_{0,i}^s$, se tendrá entonces que:

$$\Delta_2 \rightarrow 0$$

Elemento \bar{T}_i :

Consideremos la descomposición

$$\bar{T}_i - T_{0,i} = \bar{T}_i - E[\bar{T}_i] + E[\bar{T}_i] - T_{0,i} = \Delta_1 + \Delta_2$$

Donde Δ_1 es la parte aleatoria y Δ_2 la parte determinística.

$$\begin{aligned} \Delta_2 &= \left\{ \frac{1}{TM} \sum_{t=1}^T \sum_{m=1}^M T_{i,tm} \right\} - T_{0,i} = \\ &= E \left\{ \frac{1}{TM} \sum_{t=1}^T \sum_{m=1}^M \int x^i Y_{tm} \delta(x, X_{tm}) dx \right\} - T_{0,i} = E \left\{ \int x^i Y \delta(x, X) dx \right\} - T_{0,i} \end{aligned}$$

Entonces:

$$E \left\{ \int x^i Y \delta(x, X) dx \right\} = E \left\{ E \left[\left(\int x^i Y \delta(x, X) dx \right) / X \right] \right\} =$$

$$E \left\{ \left(\int x^i \delta(x, X) dx \right) E \left[Y / X \right] \right\} = E \left\{ \left(\int x^i \delta(x, X) dx \right) \alpha(X) \right\}$$

Teniendo en cuenta que $\alpha(X) = \sum_{h=1}^q \sum_{t=1}^M \theta_h I'(X) X^h + \sum_{q=h+1}^p \theta_h X^h$

$$\begin{aligned} E \left\{ \left(\int x^i \delta(x, X) dx \right) \alpha(X) \right\} &= E \left\{ \left(\int x^i \delta(x, X) dx \right) \left(\sum_{h=1}^q \sum_{t=1}^M \theta_h I'(X) X^h + \sum_{q=h+1}^p \theta_h X^h \right) \right\} = \\ &= \int \left(\int x^i \delta(x, u) dx \right) \left(\sum_{h=1}^q \sum_{t=1}^M \theta_h I'(u) u^h + \sum_{q=h+1}^p \theta_h u^h \right) f(u) du \end{aligned}$$

Usando la **proposición (2.2)** se deduce, por el *teorema de la convergencia dominada*, que:

$$\begin{aligned}
 E\left\{\left(\int x^i \delta(x, X) dx\right) \alpha(X)\right\} &\rightarrow \int u^i \left(\sum_{h=1}^q \sum_{l=1}^M \theta_h' I'(u) u^h + \sum_{q=h+1}^p \theta_h u^h \right) f(u) du = \\
 &= \sum_{h=1}^q \sum_{l=1}^M \theta_h' \left(\int u^i I'(u) u^h f(u) du \right) + \sum_{q=h+1}^p \theta_h \left(\int u^i u^h f(u) du \right) = \\
 &= \sum_{h=1}^q \sum_{l=1}^M \theta_h' E[x^i I'(x) x^h] + \sum_{q=h+1}^p \theta_h E[x^i x^h] = \\
 &= \sum_{h=1}^q \sum_{l=1}^M \theta_h' E[I'(x) x^i x^h] + \sum_{q=h+1}^p \theta_h E[x^i x^h] = \bar{T}_i
 \end{aligned}$$

Con esto se ve que $E\left\{\int x^i Y \delta(x, X) dx\right\} \rightarrow T_{0,i}$, se tendrá entonces que:

$$\Delta_2 \rightarrow 0$$

Elemento \bar{V}_{ij}^r :

Consideremos la descomposición

$$\bar{V}_{ij}^r - V_{0,ij}^r = \bar{V}_{ij}^r - E[\bar{V}_{ij}^r] + E[\bar{V}_{ij}^r] - V_{0,ij}^r = \Delta_1 + \Delta_2$$

Donde Δ_1 es la parte aleatoria y Δ_2 la parte determinística.

$$\begin{aligned} \Delta_2 &= \left\{ \frac{1}{TM} \sum_{t=1}^T \sum_{m=1}^M V_{ij,tm}^r \right\} - V_{0,ij}^r = \\ &= E \left\{ \frac{1}{TM} \sum_{t=1}^T \sum_{m=1}^M \int x^i I'(x) x^j \delta(x, X_{tm}) dx \right\} - V_{0,ij}^r = E \left\{ \int x^i I'(x) x^j \delta(x, X) dx \right\} - V_{0,ij}^r \end{aligned}$$

Entonces:

$$E \left\{ \int x^i I'(x) x^j \delta(x, X) dx \right\} = \int \left(\int x^i I'(x) x^j \delta(x, u) dx \right) f(u) du$$

Usando la **proposición (2.2)** se deduce, por el *teorema de la convergencia dominada*, que:

$$E \left\{ \int x^i I'(x) x^j \delta(x, X) dx \right\} \rightarrow \int u^i I'(u) u^j f(u) du = E \left[I'(x) x^i x^j \right]$$

Con esto se ve que $E \left\{ \int x^i I'(x) x^j \delta(x, X) dx \right\} \rightarrow V_{0,ij}^r$, se tendrá entonces que:

$$\Delta_2 \rightarrow 0$$

Elemento \bar{V}_{ij} :

Consideremos la descomposición

$$\bar{V}_{ij} - V_{0,ij} = \bar{V}_{ij} - E[\bar{V}_{ij}] + E[\bar{V}_{ij}] - V_{0,ij} = \Delta_1 + \Delta_2$$

Donde Δ_1 es la parte aleatoria y Δ_2 la parte determinística.

$$\begin{aligned} \Delta_2 &= \left\{ \frac{1}{TM} \sum_{t=1}^T \sum_{m=1}^M V_{ij,tm} \right\} - V_{0,ij} = \\ &= E \left\{ \frac{1}{TM} \sum_{t=1}^T \sum_{m=1}^M \int x^i x^j \delta(x, X_{tm}) dx \right\} - V_{0,ij} = E \left\{ \int x^i x^j \delta(x, X) dx \right\} - V_{0,ij} \end{aligned}$$

Entonces:

$$E \left\{ \int x^i x^j \delta(x, X) dx \right\} = \int \left(\int x^i x^j \delta(x, u) dx \right) f(u) du$$

Usando la **proposición (2.2)** se deduce, por el *teorema de la convergencia dominada*,

que:

$$E \left\{ \int x^i x^j \delta(x, X) dx \right\} \rightarrow \int u^i u^j f(u) du = E[x^i x^j]$$

Con esto se ve que $E \left\{ \int x^i x^j \delta(x, X) dx \right\} \rightarrow V_{0,ij}$, se tendrá entonces que:

$$\Delta_2 \rightarrow 0$$

Para demostrar que $\Delta_1 \rightarrow 0$ c.s. se emplearía la *desigualdad de Bennet*, aplicada a las variables:

$$R_h = T_{i,h} - E[T_{i,h}], R_h = T'_{i,h} - E[T'_{i,h}], R_h = V_{ij,h} - E[V_{ij,h}], R_h = V'_{ij,h} - E[V'_{ij,h}]$$

con $h=1, \dots, n$; siendo $n=TM$.

Los R_h están acotados con probabilidad 1, ya que Y está acotado y el soporte de f esta contenido en un conjunto compacto S .

$E[R_h] = 0$, ya que es una esperanza de la forma $E[V-E(V)]$ o $E[T-E(T)]$.

$E[R_h^2] < \infty$, ya que es de la forma $E[(V-E(V))^2] = \sigma_v^2$ o $E[(T-E(T))^2] = \sigma_T^2$.

Aplicando la desigualdad se tiene que $P\{|\bar{R}| \geq \varepsilon\} \leq 2e^{-n\varepsilon^2 o(1)}$.

Para tener la convergencia casi segura necesitamos que:

$$\lim_{n \rightarrow \infty} P\{|\bar{R}| < \varepsilon\} = 1$$

Aplicando el *lema de Borel-Cantelli* se tiene que $\overline{\lim}_{n \rightarrow \infty} P\{|\bar{R}| \geq \varepsilon\} = 0$, ya que

$$\sum_{n=1}^{\infty} P\{|\bar{R}| \geq \varepsilon\} \leq \sum_{n=1}^{\infty} 2e^{-n\varepsilon^2 o(1)} < \infty$$

Con esto se tiene que $\lim_{n \rightarrow \infty} P\{|\bar{R}| \geq \varepsilon\} = 0$ y por lo tanto:

$$\lim_{n \rightarrow \infty} P\{|\bar{R}| < \varepsilon\} = 1$$

y se tiene la convergencia casi segura para Δ_1 .

(apartado 2):

Considero la descomposición

$$\sqrt{n} \cdot \begin{pmatrix} \bar{V}_{ij}^r - V_{0,ij}^r \\ \bar{V}_{ij} - V_{0,ij} \\ \bar{T}_k^s - T_{0,k}^s \\ \bar{T}_k - T_{0,k} \end{pmatrix} = \sqrt{n} \cdot \begin{pmatrix} \bar{V}_{ij}^r - E[\bar{V}_{ij}^r] \\ \bar{V}_{ij} - E[\bar{V}_{ij}] \\ \bar{T}_k^s - E[\bar{T}_k^s] \\ \bar{T}_k - E[\bar{T}_k] \end{pmatrix} - \begin{pmatrix} E[\bar{V}_{ij}^r] - V_{0,ij}^r \\ E[\bar{V}_{ij}] - V_{0,ij} \\ E[\bar{T}_k^s] - T_{0,k}^s \\ E[\bar{T}_k] - T_{0,k} \end{pmatrix} = \Delta_1 + \Delta_2$$

Por hipótesis del **apartado (2)** $\Delta_2 \rightarrow 0$, y sólo necesitamos demostrar que $\Delta_2 \rightarrow N_c(0, \mathcal{A})$

Aquí, sólo necesitamos aplicar el *criterio de Liapunov generalizado* al vector

$$H_{nh} = \sqrt{n} \cdot \begin{pmatrix} \frac{V_{jh}^r - E[V_{jh}^r]}{n} \\ \frac{V_{jh} - E[V_{jh}]}{n} \\ \frac{T_{kh}^s - E[T_{kh}^s]}{n} \\ \frac{T_{kh} - E[T_{kh}]}{n} \end{pmatrix}, h = 1, 2, \dots, n; \quad n = T \cdot M.$$

Es necesario probar:

$$\sum_{i=1}^n E\{H_{ni}^h H_{ni}^k\} \rightarrow a_{nk}; \quad \text{e que: } \sum_{i=1}^n \int_{\{\|H_{ni}\| > \varepsilon\}} \|H_{ni}\|^2 dP \rightarrow 0 \quad \forall \varepsilon > 0$$

donde H_{ni}^k representa la k -ésima coordenada del vector H_{ni} , a_{nk} es el elemento (h, k)

de la matriz \mathcal{A} , y P es la medida de probabilidad inducida por el vector H_{ni} .

La segunda condición se obtiene de las siguientes relaciones:

$$\begin{aligned} \sum_{h=1}^n \int_{\{\|H_{nh}\| > \varepsilon\}} \|H_{nh}\|^2 dP &= \sum_{h=1}^n \int_{\{\|H_{nh}/\varepsilon\| > 1\}} \varepsilon^2 \left\| \frac{H_{nh}}{\varepsilon} \right\|^2 dP \leq \\ &\leq \sum_{h=1}^n \int_{\{\|H_{nh}/\varepsilon\| > 1\}} \varepsilon^2 \left\| \frac{H_{nh}}{\varepsilon} \right\|^{2+\gamma} dP \leq \frac{1}{\varepsilon^\gamma} \sum_{h=1}^n E \left\{ \|H_{nh}\|^{2+\gamma} \right\} = \\ &= \frac{1}{\varepsilon^\gamma} \sum_{h=1}^n E \left\{ \left[\dots \left(\frac{V'_{y,h} - E[V'_{y,h}]}{n} \right)^2 + \dots + \left(\frac{V'_{y,h} - E[V'_{y,h}]}{n} \right)^2 + \dots + \left(\frac{T'_{k,h} - E[T'_{k,h}]}{n} \right)^2 + \dots + \left(\frac{T'_{k,h} - E[T'_{k,h}]}{n} \right)^2 + \dots \right]^{\frac{(2+\gamma)}{2}} \right\} \approx n^{-\frac{\gamma}{2}} \rightarrow 0 \end{aligned}$$

Voy a calcular ahora los elementos de la matriz \mathcal{A} :

Elemento $(\bar{V}_{ij}^r, \bar{V}_{kl}^s)$:

$$\begin{aligned} \text{En este caso } \sum_{h=1}^n E \{ H_{nh}^x H_{nh}^y \} &= \frac{1}{n} \sum_{h=1}^n E \left\{ \left(V'_{y,h} - E[V'_{y,h}] \right) \left(V^s_{kl,h} - E[V^s_{kl,h}] \right) \right\} = \\ &= E \left\{ \left(V'_{y,h} - E[V'_{y,h}] \right) \left(V^s_{kl,h} - E[V^s_{kl,h}] \right) \right\} = \text{Cov}(V_{y,h}^r, V_{kl,h}^s). \end{aligned}$$

En general, todos los elementos a_{nk} de la matriz \mathcal{A} pueden ser obtenidos mediante el calculo de la covarianza de los elementos que lo determinan.

$$\begin{aligned} \text{Cov}(V_{y,h}^r, V_{kl,h}^s) &= E \left\{ \left(\int x^j x^j I^r(x) \delta(x, X) dx \right) \left(\int x^k x^k I^s(x) \delta(x, X) dx \right) \right\} - \\ &- E \left\{ \left(\int x^j x^j I^r(x) \delta(x, X) dx \right) \right\} E \left\{ \left(\int x^k x^k I^s(x) \delta(x, X) dx \right) \right\} = \\ &= \int \left(\int x^j x^j I^r(x) \delta(x, u) dx \right) \left(\int x^k x^k I^s(x) \delta(x, u) dx \right) f(u) du - \\ &- \left[\int \left(\int x^j x^j I^r(x) \delta(x, u) dx \right) f(u) du \right] \left[\int \left(\int x^k x^k I^s(x) \delta(x, u) dx \right) f(u) du \right] \end{aligned}$$

Teniendo en cuenta la **proposición (2.2)**, y que además existe el cuarto momento de $\|X\|$, podemos aplicar el *th de convergencia dominada* y entonces:

$$\begin{aligned} \text{Cov}(V_{ij}^r, V_{kl}^s) &\rightarrow \int u^i u^j I^r(u) u^k u^l I^s(u) f(u) du - \\ &\quad - \left[\int u^i u^j I^r(u) f(u) du \right] \left[\int u^k u^l I^s(u) f(u) du \right] = \\ &= E[x^i x^j I^r(x) x^k x^l I^s(x)] - E[x^i x^j I^r(x)] \cdot E[x^k x^l I^s(x)] = \text{Cov}(I^r(x) x^i x^j, I^s(x) x^k x^l) \end{aligned}$$

Entonces el elemento $(\bar{V}_{ij}^r, \bar{V}_{kl}^s)$ será:

$$\text{Si } r=s \quad \text{Cov}(I^r(x) x^i x^j, I^r(x) x^k x^l)$$

$$\text{Si } r \neq s \quad 0$$

Elemento $(\bar{V}_{ij}^r, \bar{V}_{kl}^s)$:

$$\begin{aligned} \text{Cov}(V_{ij}^r, V_{kl}^s) &= E\left\{\left(\int x^i x^j I^r(x) \delta(x, X) dx\right) \left(\int x^k x^l I^s(x) \delta(x, X) dx\right)\right\} - \\ &\quad - E\left\{\left(\int x^i x^j I^r(x) \delta(x, X) dx\right)\right\} E\left\{\left(\int x^k x^l I^s(x) \delta(x, X) dx\right)\right\} = \\ &= \int \left(\int x^i x^j I^r(x) \delta(x, u) dx\right) \left(\int x^k x^l I^s(x) \delta(x, u) dx\right) f(u) du - \\ &\quad - \left[\int \left(\int x^i x^j I^r(x) \delta(x, u) dx\right) f(u) du\right] \left[\int \left(\int x^k x^l I^s(x) \delta(x, u) dx\right) f(u) du\right] \end{aligned}$$

Teniendo en cuenta la **proposición (2.2)**, y que además existe el cuarto momento de $\|X\|$, podemos aplicar el *th de convergencia dominada* y entonces:

$$\begin{aligned} \text{Cov}(V'_{ij}, V_{kl}) &\rightarrow \int u^i u^j I'(u) u^k u^l f(u) du - \\ &\quad - \left[\int u^i u^j I'(u) f(u) du \right] \left[\int u^k u^l f(u) du \right] = \\ &= E[x^i x^j I'(x) x^k x^l] - E[x^i x^j I'(x)] \cdot E[x^k x^l] = \text{Cov}(I'(x) x^i x^j, x^k x^l) \end{aligned}$$

Entonces el elemento $(\bar{V}'_{ij}, \bar{V}_{kl})$ será:

$$\text{Cov}(I'(x) x^i x^j, x^k x^l)$$

Elemento $(\bar{V}'_{ij}, \bar{T}_k^s)$:

$$\begin{aligned} \text{Cov}(V'_{ij}, T_k^s) &= E\left\{\left(\int x^i x^j I'(x) \delta(x, X) dx\right) \left(\int x^k I^s(x) Y \delta(x, X) dx\right)\right\} - \\ &\quad - E\left\{\left(\int x^i x^j I'(x) \delta(x, X) dx\right)\right\} E\left\{\left(\int x^k I^s(x) Y \delta(x, X) dx\right)\right\} = \\ &= E\left\{E\left[\left(\int x^i x^j I'(x) \delta(x, X) dx\right) \left(\int x^k I^s(x) Y \delta(x, X) dx\right) \middle| X\right]\right\} - \\ &\quad - E\left\{E\left[\left(\int x^i x^j I'(x) \delta(x, X) dx\right) \middle| X\right]\right\} E\left\{E\left[\left(\int x^k I^s(x) Y \delta(x, X) dx\right) \middle| X\right]\right\} = \\ &= E\left\{\left(\int x^i x^j I'(x) \delta(x, X) dx\right) \left(\int x^k I^s(x) \delta(x, X) dx\right) E\left[Y \middle| X\right]\right\} - \\ &\quad - E\left\{\left(\int x^i x^j I'(x) \delta(x, X) dx\right)\right\} E\left\{\left(\int x^k I^s(x) \delta(x, X) dx\right) E\left[Y \middle| X\right]\right\} = \end{aligned}$$

$$\begin{aligned}
 &= E\left\{\left(\int x^i x^j I^r(x) \delta(x, X) dx\right) \left(\int x^k I^s(x) \delta(x, X) dx\right) \alpha(X)\right\} - \\
 &\quad - E\left\{\left(\int x^i x^j I^r(x) \delta(x, X) dx\right)\right\} E\left\{\left(\int x^k I^s(x) \delta(x, X) dx\right) \alpha(X)\right\} = \\
 &= \int \left(\int x^i x^j I^r(x) \delta(x, u) dx\right) \left(\int x^k I^s(x) \delta(x, u) dx\right) \alpha(u) f(u) du - \\
 &\quad - \left[\int \left(\int x^i x^j I^r(x) \delta(x, u) dx\right) f(u) du\right] \left[\int \left(\int x^k I^s(x) \delta(x, u) dx\right) \alpha(u) f(u) du\right]
 \end{aligned}$$

Teniendo en cuenta la **proposición (2.2)**, y que además existe el cuarto momento de $\|X\|$, podemos aplicar el *th de convergencia dominada* y entonces:

$$\begin{aligned}
 \text{Cov}(V_{ij}^r, T_k^s) &\rightarrow \int u^i u^j I^r(u) u^k I^s(u) \alpha(u) f(u) du - \\
 &\quad - \left[\int u^i u^j I^r(u) f(u) du\right] \left[\int u^k I^s(u) \alpha(u) f(u) du\right] = \\
 &= E\left[x^i x^j I^r(x) x^k I^s(x) \alpha(X)\right] - E\left[x^i x^j I^r(x)\right] \cdot E\left[x^k I^s(x) \alpha(X)\right]
 \end{aligned}$$

Teniendo en cuenta la expresión de $\alpha(X)$ y las propiedades de la esperanza:

$$\begin{aligned}
 \text{Cov}(V_{ij}^r, T_k^s) &\rightarrow \sum_{i=1}^M \sum_{h=1}^q \theta_h E\left[x^i x^j I^r(x) x^k I^s(x) I'(x) x^h\right] + \sum_{h=q+1}^p \theta_h E\left[x^i x^j I^r(x) x^k I^s(x) x^h\right] - \\
 &\quad - \left(\sum_{i=1}^M \sum_{h=1}^q \theta_h E\left[x^i x^j I^r(x)\right] E\left[x^k I^s(x) I'(x) x^h\right] + \sum_{h=q+1}^p \theta_h E\left[x^i x^j I^r(x)\right] E\left[x^k I^s(x) x^h\right]\right) =
 \end{aligned}$$

$$= \sum_{h=1}^q \theta_h E[x^i x^j I'(x) x^k I^s(x) x^h] + \sum_{h=q+1}^p \theta_h E[x^i x^j I'(x) x^k I^s(x) x^h] -$$

$$- \left(\sum_{i=1}^M \sum_{h=1}^q \theta_h E[x^i x^j I'(x)] E[x^k I^s(x) x^h] + \sum_{h=q+1}^p \theta_h E[x^i x^j I'(x)] E[x^k I^s(x) x^h] \right)$$

Entonces el elemento $(\bar{V}_{ij}^r, \bar{T}_k^s)$ será:

$$\text{Si } r=s \quad \sum_{h=1}^q \theta_h \text{Cov}(I'(x) x^i x^j, I'(x) x^k x^h) + \sum_{h=q+1}^p \theta_h \text{Cov}(I'(x) x^i x^j, I'(x) x^k x^h)$$

Si $r \neq s$

0

Elemento $(\bar{V}_{ij}^r, \bar{T}_k^s)$:

$$\text{Cov}(\bar{V}_{ij}^r, \bar{T}_k^s) = E\left\{\left(\int x^i x^j I'(x) \delta(x, X) dx\right) \left(\int x^k Y \delta(x, X) dx\right)\right\} -$$

$$E\left\{\left(\int x^i x^j I'(x) \delta(x, X) dx\right)\right\} E\left\{\left(\int x^k Y \delta(x, X) dx\right)\right\} =$$

$$= E\left\{E\left[\left(\int x^i x^j I'(x) \delta(x, X) dx\right) \left(\int x^k Y \delta(x, X) dx\right) \middle/ X\right]\right\} -$$

$$- E\left\{E\left[\left(\int x^i x^j I'(x) \delta(x, X) dx\right) \middle/ X\right]\right\} E\left\{E\left[\left(\int x^k Y \delta(x, X) dx\right) \middle/ X\right]\right\} =$$

$$= E\left\{\left(\int x^i x^j I'(x) \delta(x, X) dx\right) \left(\int x^k \delta(x, X) dx\right) E\left[Y \middle/ X\right]\right\} -$$

$$- E\left\{\left(\int x^i x^j I'(x) \delta(x, X) dx\right)\right\} E\left\{\left(\int x^k \delta(x, X) dx\right) E\left[Y \middle/ X\right]\right\} =$$

$$\begin{aligned}
&= E\left\{\left(\int x^i x^j I^r(x) \delta(x, X) dx\right) \left(\int x^k \delta(x, X) dx\right) \alpha(X)\right\} - \\
&\quad - E\left\{\left(\int x^i x^j I^r(x) \delta(x, X) dx\right)\right\} E\left\{\left(\int x^k \delta(x, X) dx\right) \alpha(X)\right\} = \\
&= \int \left(\int x^i x^j I^r(x) \delta(x, u) dx\right) \left(\int x^k \delta(x, u) dx\right) \alpha(u) f(u) du - \\
&\quad - \left[\int \left(\int x^i x^j I^r(x) \delta(x, u) dx\right) f(u) du\right] \left[\int \left(\int x^k \delta(x, u) dx\right) \alpha(u) f(u) du\right]
\end{aligned}$$

Teniendo en cuenta la **proposición (2.2)**, y que además existe el cuarto momento de $\|X\|$, podemos aplicar el *th de convergencia dominada* y entonces:

$$\begin{aligned}
\text{Cov}(V'_{ij}, T_k) &\rightarrow \int u^i u^j I^r(u) u^k \alpha(u) f(u) du - \left[\int u^i u^j I^r(u) f(u) du \right] \left[\int u^k \alpha(u) f(u) du \right] = \\
&= E[x^i x^j I^r(x) x^k \alpha(X)] - E[x^i x^j I^r(x)] \cdot E[x^k \alpha(X)]
\end{aligned}$$

Teniendo en cuenta la expresión de $\alpha(X)$ y las propiedades de la esperanza:

$$\begin{aligned}
\text{Cov}(V'_{ij}, T_k) &\rightarrow \sum_{i=1}^M \sum_{h=1}^q \theta_h E[x^i x^j I^r(x) x^k I^r(x) x^h] + \sum_{h=q+1}^p \theta_h E[x^i x^j I^r(x) x^k x^h] - \\
&\quad - \left(\sum_{i=1}^M \sum_{h=1}^q \theta_h E[x^i x^j I^r(x)] E[x^k I^r(x) x^h] + \sum_{h=q+1}^p \theta_h E[x^i x^j I^r(x)] E[x^k x^h] \right) = \\
&= \sum_{h=1}^q \theta_h E[x^i x^j I^r(x) x^k x^h] + \sum_{h=q+1}^p \theta_h E[x^i x^j I^r(x) x^k x^h] - \\
&\quad - \left(\sum_{h=1}^q \theta_h E[x^i x^j I^r(x)] E[x^k x^h] + \sum_{h=q+1}^p \theta_h E[x^i x^j I^r(x)] E[x^k x^h] \right) =
\end{aligned}$$

Entonces el elemento $(\bar{V}_{ij}^r, \bar{T}_k)$ será:

$$\sum_{h=1}^q \theta_h \text{Cov}(I^r(x) x^i x^j, x^k x^h) + \sum_{h=q+1}^p \theta_h \text{Cov}(I^r(x) x^i x^j, x^k x^h)$$

Elemento $(\bar{V}_{ij}, \bar{V}_{kl})$:

$$\begin{aligned} \text{Cov}(V_{ij}, V_{kl}) &= E\left\{\left(\int x^i x^j \delta(x, X) dx\right)\left(\int x^k x^l \delta(x, X) dx\right)\right\} - \\ &\quad - E\left\{\left(\int x^i x^j \delta(x, X) dx\right)\right\} E\left\{\left(\int x^k x^l \delta(x, X) dx\right)\right\} = \\ &= \int \left(\int x^i x^j \delta(x, u) dx\right) \left(\int x^k x^l \delta(x, u) dx\right) f(u) du - \\ &\quad - \left[\int \left(\int x^i x^j \delta(x, u) dx\right) f(u) du\right] \left[\int \left(\int x^k x^l \delta(x, u) dx\right) f(u) du\right] \end{aligned}$$

Teniendo en cuenta la **proposición (2.2)**, y que además existe el cuarto momento de $\|X\|$, podemos aplicar el *th de convergencia dominada* y entonces:

$$\begin{aligned} \text{Cov}(V_{ij}, V_{kl}) &\rightarrow \int u^i u^j u^k u^l f(u) du - \left[\int u^i u^j f(u) du\right] \left[\int u^k u^l f(u) du\right] = \\ &= E[x^i x^j x^k x^l] - E[x^i x^j I^r(x)] \cdot E[x^k x^l] = \text{Cov}(x^i x^j, x^k x^h) \end{aligned}$$

Entonces el elemento $(\bar{V}_{ij}, \bar{V}_{kl})$ será:

$$\text{Cov}(x^i x^j, x^k x^h)$$

Elemento $(\bar{V}_{ij}, \bar{T}_k^s)$:

$$\begin{aligned}
 \text{Cov}(V_{ij}, T_k^s) &= E\left\{\left(\int x^i x^j \delta(x, X) dx\right)\left(\int x^k I^s(x) Y \delta(x, X) dx\right)\right\} - \\
 &E\left\{\left(\int x^i x^j \delta(x, X) dx\right)\right\} E\left\{\left(\int x^k I^s(x) Y \delta(x, X) dx\right)\right\} = \\
 &= E\left\{E\left[\frac{\left(\int x^i x^j \delta(x, X) dx\right)\left(\int x^k I^s(x) Y \delta(x, X) dx\right)}{X}\right]\right\} - \\
 &- E\left\{E\left[\frac{\left(\int x^i x^j \delta(x, X) dx\right)}{X}\right] E\left[\frac{\left(\int x^k I^s(x) Y \delta(x, X) dx\right)}{X}\right]\right\} = \\
 &= E\left\{\left(\int x^i x^j \delta(x, X) dx\right)\left(\int x^k I^s(x) \delta(x, X) dx\right) E\left[\frac{Y}{X}\right]\right\} - \\
 &- E\left\{\left(\int x^i x^j \delta(x, X) dx\right)\right\} E\left\{\left(\int x^k I^s(x) \delta(x, X) dx\right) E\left[\frac{Y}{X}\right]\right\} = \\
 &= E\left\{\left(\int x^i x^j \delta(x, X) dx\right)\left(\int x^k I^s(x) \delta(x, X) dx\right) \alpha(X)\right\} - \\
 &- E\left\{\left(\int x^i x^j \delta(x, X) dx\right)\right\} E\left\{\left(\int x^k I^s(x) \delta(x, X) dx\right) \alpha(X)\right\} = \\
 &= \int \left(\int x^i x^j \delta(x, u) dx\right) \left(\int x^k I^s(x) \delta(x, u) dx\right) \alpha(u) f(u) du - \\
 &- \left[\int \left(\int x^i x^j \delta(x, u) dx\right) f(u) du\right] \left[\int \left(\int x^k I^s(x) \delta(x, u) dx\right) \alpha(u) f(u) du\right]
 \end{aligned}$$

Teniendo en cuenta la **proposición (2.2)**, y que además existe el cuarto momento de $\|X\|$, podemos aplicar el *th de convergencia dominada* y entonces:

$$\begin{aligned} \text{Cov}(V_{ij}, T_k^s) &\rightarrow \int u^i u^j u^k I^s(u) \alpha(u) f(u) du - \\ &\quad - \left[\int u^i u^j f(u) du \right] \left[\int u^k I^s(u) \alpha(u) f(u) du \right] = \\ &= E[x^i x^j x^k I^s(x) \alpha(X)] - E[x^i x^j] \cdot E[x^k I^s(x) \alpha(X)] \end{aligned}$$

Teniendo en cuenta la expresión de $\alpha(X)$ y las propiedades de la esperanza:

$$\begin{aligned} \text{Cov}(V_{ij}, T_k^s) &\rightarrow \sum_{i=1}^M \sum_{h=1}^q \theta_h E[x^i x^j x^k I^s(x) I'(x) x^h] + \sum_{h=q+1}^p \theta_h E[x^i x^j x^k I^s(x) x^h] - \\ &\quad - \left(\sum_{i=1}^M \sum_{h=1}^q \theta_h E[x^i x^j] E[x^k I^s(x) I'(x) x^h] + \sum_{h=q+1}^p \theta_h E[x^i x^j] E[x^k I^s(x) x^h] \right) = \\ &= \sum_{h=1}^q \theta_h^s E[x^i x^j x^k I^s(x) x^h] + \sum_{h=q+1}^p \theta_h E[x^i x^j x^k I^s(x) x^h] - \\ &\quad - \left(\sum_{h=1}^q \theta_h^s E[x^i x^j] E[x^k I^s(x) x^h] + \sum_{h=q+1}^p \theta_h E[x^i x^j] E[x^k I^s(x) x^h] \right) = \end{aligned}$$

Entonces el elemento $(\bar{V}_{ij}, \bar{T}_k^s)$ será:

$$\sum_{h=1}^q \theta_h^s \text{Cov}(x^i x^j, I^s(x) x^k x^h) + \sum_{h=q+1}^p \theta_h \text{Cov}(x^i x^j, I^s(x) x^k x^h)$$

Elemento $(\bar{V}_{ij}, \bar{T}_k)$:

$$\begin{aligned}
 \text{Cov}(V_{ij}, T_k) &= E\left\{\left(\int x^i x^j \delta(x, X) dx\right)\left(\int x^k Y \delta(x, X) dx\right)\right\} - \\
 &\quad E\left\{\left(\int x^i x^j \delta(x, X) dx\right)\right\} E\left\{\left(\int x^k Y \delta(x, X) dx\right)\right\} = \\
 &= E\left\{E\left[\frac{\left(\int x^i x^j \delta(x, X) dx\right)\left(\int x^k Y \delta(x, X) dx\right)}{X}\right]\right\} - \\
 &\quad - E\left\{E\left[\frac{\left(\int x^i x^j \delta(x, X) dx\right)}{X}\right]\right\} E\left\{E\left[\frac{\left(\int x^k Y \delta(x, X) dx\right)}{X}\right]\right\} = \\
 &= E\left\{\left(\int x^i x^j \delta(x, X) dx\right)\left(\int x^k \delta(x, X) dx\right) E\left[\frac{Y}{X}\right]\right\} - \\
 &\quad - E\left\{\left(\int x^i x^j \delta(x, X) dx\right)\right\} E\left\{\left(\int x^k \delta(x, X) dx\right) E\left[\frac{Y}{X}\right]\right\} = \\
 &= E\left\{\left(\int x^i x^j \delta(x, X) dx\right)\left(\int x^k \delta(x, X) dx\right) \alpha(X)\right\} - \\
 &\quad - E\left\{\left(\int x^i x^j \delta(x, X) dx\right)\right\} E\left\{\left(\int x^k \delta(x, X) dx\right) \alpha(X)\right\} = \\
 &= \int \left(\int x^i x^j \delta(x, u) dx\right) \left(\int x^k \delta(x, u) dx\right) \alpha(u) f(u) du - \\
 &\quad - \left[\int \left(\int x^i x^j \delta(x, u) dx\right) f(u) du\right] \left[\int \left(\int x^k \delta(x, u) dx\right) \alpha(u) f(u) du\right]
 \end{aligned}$$

Teniendo en cuenta la **proposición (2.2)**, y que además existe el cuarto momento de $\|X\|$, podemos aplicar el *th de convergencia dominada* y entonces:

$$\begin{aligned}\text{Cov}(V_{ij}, T_k) &\rightarrow \int u^i u^j u^k \alpha(u) f(u) du - \left[\int u^i u^j f(u) du \right] \left[\int u^k \alpha(u) f(u) du \right] = \\ &= E[x^i x^j x^k \alpha(X)] - E[x^i x^j] \cdot E[x^k \alpha(X)]\end{aligned}$$

Teniendo en cuenta la expresión de $\alpha(X)$ y las propiedades de la esperanza:

$$\begin{aligned}\text{Cov}(V_{ij}, T_k) &\rightarrow \sum_{i=1}^M \sum_{h=1}^q \theta_h E[x^i x^j x^k I'(x) x^h] + \sum_{h=q+1}^p \theta_h E[x^i x^j x^k x^h] - \\ &- \left(\sum_{i=1}^M \sum_{h=1}^q \theta_h E[x^i x^j] E[x^k I'(x) x^h] + \sum_{h=q+1}^p \theta_h E[x^i x^j] E[x^k x^h] \right)\end{aligned}$$

Entonces el elemento $(\bar{V}_{ij}, \bar{T}_k)$ será:

$$\sum_{i=1}^M \sum_{h=1}^q \theta_h \text{Cov}(x^i x^j, I'(x) x^k x^h) + \sum_{h=q+1}^p \theta_h \text{Cov}(x^i x^j, I'(x) x^k x^h)$$

Elemento $(\bar{T}_i^r, \bar{T}_k^s)$:

$$\begin{aligned}\text{Cov}(T_i^r, T_k^s) &= E\left\{\left(\int x^i I^r(x) Y \delta(x, X) dx\right) \left(\int x^k I^s(x) Y \delta(x, X) dx\right)\right\} - \\ &E\left\{\left(\int x^i I^r(x) Y \delta(x, X) dx\right)\right\} E\left\{\left(\int x^k I^s(x) Y \delta(x, X) dx\right)\right\} = \\ &= E\left\{E\left[\left(\int x^i I^r(x) Y \delta(x, X) dx\right) \left(\int x^k I^s(x) Y \delta(x, X) dx\right) \middle/ X\right]\right\} - \\ &- E\left\{E\left[\left(\int x^i I^r(x) Y \delta(x, X) dx\right) \middle/ X\right]\right\} E\left\{E\left[\left(\int x^k I^s(x) Y \delta(x, X) dx\right) \middle/ X\right]\right\} =\end{aligned}$$

$$\begin{aligned}
 &= E \left\{ \left(\int x^i I'(x) \delta(x, X) dx \right) \left(\int x^k I^s(x) \delta(x, X) dx \right) E \left[\frac{Y^2}{X} \right] \right\} - \\
 &\quad - E \left\{ \left(\int x^i I'(x) \delta(x, X) dx \right) E \left[\frac{Y}{X} \right] \right\} E \left\{ \left(\int x^k I^s(x) \delta(x, X) dx \right) E \left[\frac{Y}{X} \right] \right\} = \\
 &= E \left\{ \left(\int x^i I'(x) \delta(x, X) dx \right) \left(\int x^k I^s(x) \delta(x, X) dx \right) (\sigma^2 + \alpha^2(X)) \right\} - \\
 &\quad - E \left\{ \left(\int x^i I'(x) \delta(x, X) dx \right) \alpha(X) \right\} E \left\{ \left(\int x^k I^s(x) \delta(x, X) dx \right) \alpha(X) \right\} = \\
 &= \int \left(\int x^i I'(x) \delta(x, u) dx \right) \left(\int x^k I^s(x) \delta(x, u) dx \right) (\sigma^2 + \alpha^2(u)) f(u) du - \\
 &\quad - \left[\int \left(\int x^i I'(x) \delta(x, u) dx \right) \alpha(u) f(u) du \right] \left[\int \left(\int x^k I^s(x) \delta(x, u) dx \right) \alpha(u) f(u) du \right]
 \end{aligned}$$

Teniendo en cuenta la **proposición (2.2)**, y que además existe el cuarto momento de $\|X\|$, podemos aplicar el *th de convergencia dominada* y entonces:

$$\begin{aligned}
 \text{Cov}(T_i^r, T_k^s) &\rightarrow \int u^i I'(u) u^k I^s(u) (\sigma^2 + \alpha^2(u)) f(u) du - \\
 &\quad - \left[\int u^i I'(u) \alpha(u) f(u) du \right] \left[\int u^k I^s(u) \alpha(u) f(u) du \right] = \\
 &= E \left[x^i I'(x) x^k I^s(x) (\sigma^2 + \alpha^2(X)) \right] - E \left[x^i I'(x) \alpha(X) \right] \cdot E \left[x^k I^s(x) \alpha(X) \right]
 \end{aligned}$$

Teniendo en cuenta la expresión de $\alpha(X)$, tendremos que:

$$\begin{aligned}\alpha^2(x) = & \sum_{h=1}^q \sum_{t=1}^M \sum_{v=1}^q \sum_{d=1}^M \theta_h I'(X) X^h \theta_v^d I^d(X) X^v + \sum_{h=1}^q \sum_{t=1}^M \sum_{v=q+1}^p \theta_h I'(X) X^h \theta_v X^v + \\ & + \sum_{v=1}^q \sum_{d=1}^M \sum_{h=q+1}^p \theta_h X^h \theta_v^d I^d(X) X^v + \sum_{h=q+1}^p \sum_{v=q+1}^p \theta_h X^h \theta_v X^v\end{aligned}$$

y aplicando además las propiedades de la esperanza se tendrá:

$$\begin{aligned}\text{Cov}(T_i^r, T_k^s) \rightarrow & E[x^i I'(x) x^k I^s(x) \sigma^2] + \\ & + \sum_{t=1}^M \sum_{d=1}^M \sum_{h=1}^q \sum_{v=1}^q \theta_h \theta_v^d E[x^i I'(x) x^k I^s(x) I'(x) x^h I^d(x) x^v] + \\ & + \sum_{t=1}^M \sum_{h=1}^q \sum_{v=q+1}^p \theta_h \theta_v E[x^i I'(x) x^k I^s(x) I'(x) x^h x^v] + \\ & + \sum_{t=1}^M \sum_{h=q+1}^p \sum_{v=1}^q \theta_h \theta_v^d E[x^i I'(x) x^k I^s(x) x^h I^d(x) x^v] + \\ & + \sum_{h=q+1}^p \sum_{v=q+1}^p \theta_h \theta_v E[x^i I'(x) x^k I^s(x) x^h x^v] - \\ & - \left(\sum_{t=1}^M \sum_{d=1}^M \sum_{h=1}^q \sum_{v=1}^q \theta_h \theta_v^d E[x^i I'(x) I'(x) x^h] E[x^k I^s(x) I^d(x) x^v] + \right. \\ & + \sum_{t=1}^M \sum_{h=1}^q \sum_{v=q+1}^p \theta_h \theta_v E[x^i I'(x) I'(x) x^h] E[x^k I^s(x) x^v] + \\ & + \sum_{d=1}^M \sum_{h=q+1}^p \sum_{v=1}^q \theta_h \theta_v^d E[x^i I'(x) x^h] E[x^k I^s(x) I^d(x) x^v] + \\ & \left. + \sum_{h=q+1}^p \sum_{v=q+1}^p \theta_h \theta_v E[x^i I'(x) x^h] E[x^k I^s(x) x^v] \right) =\end{aligned}$$

$$\begin{aligned}
 &= E[x^i I^r(x) x^k I^s(x) \sigma^2] + \sum_{h=1}^q \sum_{v=1}^q \theta_h \theta_v E[x^i I^r(x) x^k I^s(x) x^h x^v] + \\
 &\quad + \sum_{h=1}^q \sum_{v=q+1}^p \theta_h \theta_v E[x^i I^r(x) x^k I^s(x) x^h x^v] + \sum_{h=q+1}^p \sum_{v=1}^q \theta_h \theta_v E[x^i I^r(x) x^k I^s(x) x^h x^v] + \\
 &\quad + \sum_{h=q+1}^p \sum_{v=q+1}^p \theta_h \theta_v E[x^i I^r(x) x^k I^s(x) x^h x^v] - \left(\sum_{h=1}^q \sum_{v=1}^q \theta_h \theta_v E[x^i I^r(x) x^h] E[x^k I^s(x) x^v] + \right. \\
 &\quad + \sum_{h=1}^q \sum_{v=q+1}^p \theta_h \theta_v E[x^i I^r(x) x^h] E[x^k I^s(x) x^v] + \sum_{h=q+1}^p \sum_{v=1}^q \theta_h \theta_v E[x^i I^r(x) x^h] E[x^k I^s(x) x^v] + \\
 &\quad \left. + \sum_{h=q+1}^p \sum_{v=q+1}^p \theta_h \theta_v E[x^i I^r(x) x^h] E[x^k I^s(x) x^v] \right) = \\
 &= E[x^i I^r(x) x^k I^s(x) \sigma^2] + \sum_{h=1}^q \sum_{v=1}^q \theta_h \theta_v \text{Cov}(I^r(x) x^i x^h, I^s(x) x^k x^v) + \\
 &\quad + \sum_{h=1}^q \sum_{v=q+1}^p \theta_h \theta_v \text{Cov}(I^r(x) x^i x^h, I^s(x) x^k x^v) + \\
 &\quad + \sum_{h=q+1}^p \sum_{v=1}^q \theta_h \theta_v \text{Cov}(I^r(x) x^i x^h, I^s(x) x^k x^v) + \sum_{h=q+1}^p \sum_{v=q+1}^p \theta_h \theta_v \text{Cov}(I^r(x) x^i x^h, I^s(x) x^k x^v)
 \end{aligned}$$

Entonces el elemento $(\bar{T}_i^r, \bar{T}_k^s)$ será:

Si $r=s$

$$\sum_{h=1}^q \sum_{v=1}^q \theta_h \theta_v \text{Cov}(I^r(x) x^i x^h, I^r(x) x^k x^v) + \sum_{h=1}^q \sum_{v=q+1}^p \theta_h \theta_v \text{Cov}(I^r(x) x^i x^h, I^r(x) x^k x^v) +$$

$$+ \sum_{h=q+1}^p \sum_{v=1}^q \theta_h \theta_v \text{Cov}(I'(x) x^i x^h, I'(x) x^k x^v) +$$

$$\sum_{h=q+1}^p \sum_{v=q+1}^p \theta_h \theta_v \text{Cov}(I'(x) x^i x^h, I'(x) x^k x^v) + \sigma^2 E[I'(x) x^i x^k]$$

Si $r \neq s$

0

Elemento (\bar{T}_i, \bar{T}_k) :

$$\text{Cov}(T_i, T_k) = E\left\{\left(\int x^i I'(x) Y \delta(x, X) dx\right) \left(\int x^k Y \delta(x, X) dx\right)\right\} -$$

$$E\left\{\left(\int x^i I'(x) Y \delta(x, X) dx\right)\right\} E\left\{\left(\int x^k Y \delta(x, X) dx\right)\right\} =$$

$$= E\left\{E\left[\left(\int x^i I'(x) Y \delta(x, X) dx\right) \left(\int x^k Y \delta(x, X) dx\right) \middle| X\right]\right\} -$$

$$- E\left\{E\left[\left(\int x^i I'(x) Y \delta(x, X) dx\right) \middle| X\right]\right\} E\left\{E\left[\left(\int x^k Y \delta(x, X) dx\right) \middle| X\right]\right\} =$$

$$= E\left\{\left(\int x^i I'(x) \delta(x, X) dx\right) \left(\int x^k \delta(x, X) dx\right) E\left[Y^2 \middle| X\right]\right\} -$$

$$- E\left\{\left(\int x^i I'(x) \delta(x, X) dx\right) E\left[Y \middle| X\right]\right\} E\left\{\left(\int x^k \delta(x, X) dx\right) E\left[Y \middle| X\right]\right\} =$$

$$= E\left\{\left(\int x^i I'(x) \delta(x, X) dx\right) \left(\int x^k \delta(x, X) dx\right) (\sigma^2 + \alpha^2(X))\right\} -$$

$$- E\left\{\left(\int x^i I'(x) \delta(x, X) dx\right) \alpha(X)\right\} E\left\{\left(\int x^k \delta(x, X) dx\right) \alpha(X)\right\} =$$

$$= \left(\int \int x^i I'(x) \delta(x, u) dx \right) \left(\int x^k \delta(x, u) dx \right) (\sigma^2 + \alpha^2(u)) f(u) du - \\ - \left[\left(\int \int x^i I'(x) \delta(x, u) dx \right) \alpha(u) f(u) du \right] \left[\left(\int x^k \delta(x, u) dx \right) \alpha(u) f(u) du \right]$$

Teniendo en cuenta la **proposición (2.2)**, y que además existe el cuarto momento de $\|X\|$, podemos aplicar el *th de convergencia dominada* y entonces:

$$\text{Cov}(T_i^*, T_k) \rightarrow \int u^i I'(u) u^k (\sigma^2 + \alpha^2(u)) f(u) du \\ - \left[\int u^i I'(u) \alpha(u) f(u) du \right] \left[\int u^k \alpha(u) f(u) du \right] = \\ = E \left[x^i I'(x) x^k (\sigma^2 + \alpha^2(X)) \right] - E \left[x^i I'(x) \alpha(X) \right] \cdot E \left[x^k \alpha(X) \right]$$

Teniendo en cuenta la expresión de $\alpha(X)$ y las propiedades de la esperanza se tendrá:

$$\text{Cov}(T_i^*, T_k) \rightarrow E \left[x^i I'(x) x^k \sigma^2 \right] + \sum_{i=1}^M \sum_{d=1}^M \sum_{h=1}^q \sum_{v=1}^q \theta_h \theta_v E \left[x^i I'(x) x^k I'(x) x^h I^d(x) x^v \right] + \\ + \sum_{i=1}^M \sum_{h=1}^q \sum_{v=q+1}^p \theta_h \theta_v E \left[x^i I'(x) x^k I'(x) x^h x^v \right] + \\ + \sum_{i=1}^M \sum_{h=q+1}^p \sum_{v=1}^q \theta_h \theta_v E \left[x^i I'(x) x^k x^h I^d(x) x^v \right] + \sum_{h=q+1}^p \sum_{v=q+1}^p \theta_h \theta_v E \left[x^i I'(x) x^k x^h x^v \right] - \\ - \left(\sum_{i=1}^M \sum_{d=1}^M \sum_{h=1}^q \sum_{v=1}^q \theta_h \theta_v E \left[x^i I'(x) I'(x) x^h \right] E \left[x^k I^d(x) x^v \right] + \right.$$

$$\begin{aligned}
 & + \sum_{i=1}^M \sum_{h=1}^q \sum_{v=q+1}^p \theta_h \theta_v E[x^i I'(x) I'(x) x^h] E[x^k x^v] + \\
 & + \sum_{d=1}^M \sum_{h=q+1}^p \sum_{v=1}^q \theta_h \theta_v E[x^i I'(x) x^h] E[x^k I^d(x) x^v] + \\
 & + \sum_{h=q+1}^p \sum_{v=q+1}^p \theta_h \theta_v E[x^i I'(x) x^h] E[x^k x^v] \Big) = \\
 & = E[x^i I'(x) x^k \sigma^2] + \sum_{h=1}^q \sum_{v=1}^q \theta_h \theta_v E[x^i I'(x) x^k x^h x^v] + \\
 & + \sum_{h=1}^q \sum_{v=q+1}^p \theta_h \theta_v E[x^i I'(x) x^k x^h x^v] + \sum_{h=q+1}^p \sum_{v=1}^q \theta_h \theta_v E[x^i I'(x) x^k x^h x^v] + \\
 & + \sum_{h=q+1}^p \sum_{v=q+1}^p \theta_h \theta_v E[x^i I'(x) x^k x^h x^v] - \left(\sum_{h=1}^q \sum_{v=1}^q \theta_h \theta_v E[x^i I'(x) x^h] E[x^k x^v] + \right. \\
 & + \sum_{h=1}^q \sum_{v=q+1}^p \theta_h \theta_v E[x^i I'(x) x^h] E[x^k x^v] + \sum_{h=q+1}^p \sum_{v=1}^q \theta_h \theta_v E[x^i I'(x) x^h] E[x^k x^v] + \\
 & \left. + \sum_{h=q+1}^p \sum_{v=q+1}^p \theta_h \theta_v E[x^i I'(x) x^h] E[x^k x^v] \right) = \\
 & = E[x^i I'(x) x^k \sigma^2] + \sum_{h=1}^q \sum_{v=1}^q \theta_h \theta_v \text{Cov}(I'(x) x^i x^h, x^k x^v) + \\
 & + \sum_{h=1}^q \sum_{v=q+1}^p \theta_h \theta_v \text{Cov}(I'(x) x^i x^h, x^k x^v) + \\
 & + \sum_{h=q+1}^p \sum_{v=1}^q \theta_h \theta_v \text{Cov}(I'(x) x^i x^h, x^k x^v) + \sum_{h=q+1}^p \sum_{v=q+1}^p \theta_h \theta_v \text{Cov}(I'(x) x^i x^h, x^k x^v)
 \end{aligned}$$

Entonces el elemento (\bar{T}_i, \bar{T}_k) será:

$$\begin{aligned} & \sum_{h=1}^q \sum_{v=1}^q \theta_h' \theta_v' \text{Cov}(I^r(x) x^i x^h, x^k x^v) + \sum_{h=1}^q \sum_{v=q+1}^p \theta_h' \theta_v' \text{Cov}(I^r(x) x^i x^h, x^k x^v) + \\ & + \sum_{h=q+1}^p \sum_{v=1}^q \theta_h' \theta_v' \text{Cov}(I^r(x) x^i x^h, x^k x^v) + \\ & + \sum_{h=q+1}^p \sum_{v=q+1}^p \theta_h' \theta_v' \text{Cov}(I^r(x) x^i x^h, x^k x^v) + \sigma^2 E[I^r(x) x^i x^k] \end{aligned}$$

Elemento (\bar{T}_i, \bar{T}_k) :

$$\begin{aligned} \text{Cov}(T_j, T_k) &= E\left\{\left(\int x^i Y \delta(x, X) dx\right)\left(\int x^k Y \delta(x, X) dx\right)\right\} - \\ & E\left\{\left(\int x^i Y \delta(x, X) dx\right)\right\} E\left\{\left(\int x^k Y \delta(x, X) dx\right)\right\} = \\ &= E\left\{E\left[\left(\int x^i Y \delta(x, X) dx\right)\left(\int x^k Y \delta(x, X) dx\right) \middle/ X\right]\right\} - E\left\{E\left[\left(\int x^i Y \delta(x, X) dx\right) \middle/ X\right]\right\} E\left\{E\left[\left(\int x^k Y \delta(x, X) dx\right) \middle/ X\right]\right\} = \\ &= E\left\{\left(\int x^i \delta(x, X) dx\right)\left(\int x^k \delta(x, X) dx\right) E\left[Y^2 \middle/ X\right]\right\} - \\ & - E\left\{\left(\int x^i \delta(x, X) dx\right) E\left[Y \middle/ X\right]\right\} E\left\{\left(\int x^k \delta(x, X) dx\right) E\left[Y \middle/ X\right]\right\} = \\ &= E\left\{\left(\int x^i \delta(x, X) dx\right)\left(\int x^k \delta(x, X) dx\right)\left(\sigma^2 + \alpha^2(X)\right)\right\} - \\ & - E\left\{\left(\int x^i \delta(x, X) dx\right) \alpha(X)\right\} E\left\{\left(\int x^k \delta(x, X) dx\right) \alpha(X)\right\} = \end{aligned}$$

$$= \left(\int \int x^j \delta(x, u) dx \right) \left(\int x^k \delta(x, u) dx \right) (\sigma^2 + \alpha^2(u)) f(u) du - \\ - \left[\left(\int \int x^j \delta(x, u) dx \right) \alpha(u) f(u) du \right] \left[\left(\int \int x^k \delta(x, u) dx \right) \alpha(u) f(u) du \right]$$

Teniendo en cuenta la **proposición (2.2)**, y que además existe el cuarto momento de $\|X\|$, podemos aplicar el *th de convergencia dominada* y entonces:

$$\text{Cov}(T_j, T_k) \rightarrow \int u^j u^k (\sigma^2 + \alpha^2(u)) f(u) du - \left[\int u^j \alpha(u) f(u) du \right] \left[\int u^k \alpha(u) f(u) du \right] = \\ = E[x^j x^k (\sigma^2 + \alpha^2(X))] - E[x^j \alpha(X)] E[x^k \alpha(X)]$$

Teniendo en cuenta la expresión de $\alpha(X)$ y las propiedades de la esperanza se tendrá:

$$\text{Cov}(T_j, T_k) \rightarrow E[x^j x^k \sigma^2] + \sum_{i=1}^M \sum_{d=1}^M \sum_{h=1}^q \sum_{v=1}^q \theta_h \theta_v^d E[x^j x^k I^h(x) I^d(x) x^v] + \\ + \sum_{i=1}^M \sum_{h=1}^q \sum_{v=q+1}^p \theta_h \theta_v E[x^j x^k I^h(x) x^v] + \\ + \sum_{i=1}^M \sum_{h=q+1}^p \sum_{v=1}^q \theta_h \theta_v^d E[x^j x^k x^h I^d(x) x^v] + \sum_{h=q+1}^p \sum_{v=q+1}^p \theta_h \theta_v E[x^j x^k x^h x^v] - \\ - \left(\sum_{i=1}^M \sum_{d=1}^M \sum_{h=1}^q \sum_{v=1}^q \theta_h \theta_v^d E[x^j I^h(x) x^h] E[x^k I^d(x) x^v] + \right. \\ \left. + \sum_{i=1}^M \sum_{h=1}^q \sum_{v=q+1}^p \theta_h \theta_v E[x^j I^h(x) x^h] E[x^k x^v] + \sum_{d=1}^M \sum_{h=q+1}^p \sum_{v=1}^q \theta_h \theta_v^d E[x^j x^h] E[x^k I^d(x) x^v] + \right. \\ \left. + \sum_{h=q+1}^p \sum_{v=q+1}^p \theta_h \theta_v E[x^j x^h] E[x^k x^v] \right) =$$

$$\begin{aligned}
 &= E[x^i x^k \sigma^2] + \sum_{i=1}^M \sum_{h=1}^q \sum_{v=1}^q \theta_h \theta_v E[x^i x^k I'(x) x^h x^v] + \\
 &\quad + \sum_{i=1}^M \sum_{h=1}^q \sum_{v=q+1}^p \theta_h \theta_v E[x^i x^k I'(x) x^h x^v] + \sum_{d=1}^M \sum_{h=q+1}^p \sum_{v=1}^q \theta_h \theta_v^d E[x^i x^k x^h I^d(x) x^v] + \\
 &\quad + \sum_{h=q+1}^p \sum_{v=q+1}^p \theta_h \theta_v E[x^i x^k x^h x^v] - \left(\sum_{i=1}^M \sum_{h=1}^q \sum_{v=1}^q \theta_h \theta_v E[x^i I'(x) x^h] E[x^k x^v] + \right. \\
 &\quad \left. + \sum_{h=q+1}^p \sum_{v=q+1}^p \theta_h \theta_v E[x^i x^h] E[x^k x^v] \right) = \\
 &= E[x^i x^k \sigma^2] + \sum_{i=1}^M \sum_{h=1}^q \sum_{v=q+1}^p \theta_h \theta_v \text{Cov}(I'(x) x^i x^k, x^h x^v) + \\
 &\quad + \sum_{i=1}^M \sum_{h=1}^q \sum_{v=q+1}^p \theta_h \theta_v \text{Cov}(I'(x) x^i x^k, x^h x^v) + \\
 &\quad + \sum_{i=1}^M \sum_{h=q+1}^p \sum_{v=1}^q \theta_h \theta_v \text{Cov}(I'(x) x^i x^h, x^k x^v) + \sum_{h=q+1}^p \sum_{v=q+1}^p \theta_h \theta_v \text{Cov}(x^i x^h, x^k x^v)
 \end{aligned}$$

Entonces el elemento (\bar{T}_i, \bar{T}_k) será:

$$\begin{aligned}
 &\sum_{h=1}^q \sum_{v=1}^q \sum_{i=1}^M \theta_h \theta_v \text{Cov}(I'(x) x^i x^h, x^k x^v) + \sum_{h=1}^q \sum_{v=q+1}^p \sum_{i=1}^M \theta_h \theta_v \text{Cov}(I'(x) x^i x^h, x^k x^v) + \\
 &\quad + \sum_{h=q+1}^p \sum_{v=1}^q \sum_{i=1}^M \theta_h \theta_v \text{Cov}(I'(x) x^i x^h, x^k x^v) + \sum_{h=q+1}^p \sum_{v=q+1}^p \theta_h \theta_v \text{Cov}(x^i x^h, x^k x^v) + \sigma^2 E[x^i x^k]
 \end{aligned}$$

Con esto queda completada la demostración de este lema.

Demostración del teorema (2.4)**(apartado 1)**

Tenemos que demostrar la convergencia casi segura del estimador $\hat{\Theta}$ al parámetro Θ , i.e.:

$$\hat{\Theta} = C^{-1} D \xrightarrow{c.s.} \Theta$$

Los componentes de $\hat{\Theta}$ están contruidos mediante combinaciones de elementos de la forma $\bar{V}_{ik}^r, \bar{V}_{ik}, \bar{T}_k^s, \bar{T}_k$. Como por el lema (2.3) estos elementos convergen casi seguro, las combinaciones que usamos para construir $\hat{\Theta}$ también convergerán casi seguro al valor dado por la sustitución en esas combinaciones de los límites de los valores $\bar{V}_{ik}^r, \bar{V}_{ik}, \bar{T}_k^s, \bar{T}_k$.

Tendremos por lo tanto que:

$$\hat{\Theta} = C^{-1} D \xrightarrow{c.s.} C_0^{-1} D_0$$

Siendo C_0 la matriz que se obtiene de sustituir los límites de los elementos $\bar{V}_{ik}^r, \bar{V}_{ik}$, (o sea $\bar{V}_{ik}^r, \bar{V}_{ik}$), en la matriz C , y D_0 de sustituir los límites de los elementos \bar{T}_k^s, \bar{T}_k , (o sea $T_{0,k}^r, T_{0,k}$) en la matriz D .

Ahora sólo queda demostrar que $C_0^{-1} D_0 = \Theta$.

Si denomino C_0^* a la matriz de adjuntos de C_0 , tengo que $C_0^{-1} = \frac{C_0^*}{|C_0|}$.

La matriz C_0^* tomará la forma:

$$\begin{pmatrix} W_0^{1,11} & W_0^{1,1q} & B_0^{1,1/2,1} & B_0^{1,1/2,q} & B_0^{1,1/M,1} & B_0^{1,1/M,q} & W_0^{1,1q+1} & W_0^{1,1p} \\ W_0^{1,q1} & W_0^{1,qq} & B_0^{1,q/2,1} & B_0^{1,q/2,q} & B_0^{1,q/M,1} & B_0^{1,q/M,q} & W_0^{1,qq+1} & W_0^{1,qp} \\ B_0^{2,1/1,1} & B_0^{2,1/2,q} & W_0^{2,11} & W_0^{2,1q} & B_0^{2,1/M,1} & B_0^{2,1/M,q} & W_0^{2,1q+1} & W_0^{2,1p} \\ B_0^{2,q/1,1} & B_0^{2,q/1,q} & W_0^{2,q1} & W_0^{2,qq} & B_0^{2,q/M,1} & B_0^{2,q/M,q} & W_0^{2,qq+1} & W_0^{2,qp} \\ B_0^{M,1/1,1} & B_0^{M,1/1,q} & B_0^{M,1/2,1} & B_0^{M,1/2,q} & W_0^{M,11} & W_0^{M,1q} & W_0^{M,1q+1} & W_0^{M,1p} \\ B_0^{M,q/1,1} & B_0^{M,q/1,q} & B_0^{M,q/2,1} & B_0^{M,q/2,q} & W_0^{M,q1} & W_0^{M,qq} & W_0^{M,qq+1} & W_0^{M,qp} \\ W_0^{1,q+11} & W_0^{1,q+1q} & W_0^{2,q+11} & W_0^{2,q+1q} & W_0^{M,q+11} & W_0^{M,q+1q} & W_0^{q+1q+1} & W_0^{q+1p} \\ W_0^{1,p1} & W_0^{1,pq} & W_0^{2,p1} & W_0^{2,pq} & W_0^{M,p1} & W_0^{M,pq} & W_0^{pq+1} & W_0^{pp} \end{pmatrix}$$

Siendo los elementos de la forma $W_0^{r,ij}$ los adjuntos de elementos de C_0 de la forma

$V_{0,ij}^r$, y los elementos $B_0^{r,i,j}$ adjuntos de un cero de C_0 , situado en la misma fila que

los elementos $V_{0,ik}^r$, $k=1,2,\dots,p$, y en la misma columna que los elementos $V_{0,jl}^s$,

$l=1,2,\dots,p$, con lo que estos elementos faltarán del adjunto.

Para la construcción de D_0 tendremos en cuenta las formas de los elementos $V_{0,ik}^r$,

$V_{0,ik}^r$, $T_{0,k}^r$, $T_{0,k}$, que según el lema (2.3) son:

$$V_{0,ij}^r = E[I^r x^i x^j], \quad V_{0,ij} = E[x^i x^j],$$

$$T_{0,k}^s = \sum_{h=1}^q \theta_h^s \cdot E[I^s x^k x^h] + \sum_{h=q+1}^p \theta_h \cdot E[I^s x^k x^h]$$

$$T_{0,k} = \sum_{h=1}^q \sum_{l=1}^M \theta_h^l \cdot E[I^l x^k x^h] + \sum_{h=q+1}^p \theta_h \cdot E[x^k x^h]$$

De lo que se deduce:

$$T_{0,k}^s = \sum_{h=1}^q \theta_h^s \cdot V_{0,kh}^s + \sum_{h=q+1}^p \theta_h \cdot V_{0,kh}^s; \quad T_{0,k} = \sum_{h=1}^q \sum_{t=1}^M \theta_h^t \cdot V_{0,kh}^t + \sum_{h=q+1}^p \theta_h \cdot V_{0,kh}$$

Por lo tanto:

$$D_0 = \begin{pmatrix} T_{0,1}^1 \\ T_{0,q}^1 \\ T_{0,1}^2 \\ \vdots \\ T_{0,q}^M \\ T_{0,q+1} \\ \vdots \\ T_{0,p} \end{pmatrix} = \begin{pmatrix} \sum_{h=1}^q \theta_h^1 \cdot V_{0,1h}^1 + \sum_{h=q+1}^p \theta_h \cdot V_{0,1h}^1 \\ \sum_{h=1}^q \theta_h^1 \cdot V_{0,qh}^1 + \sum_{h=q+1}^p \theta_h \cdot V_{0,qh}^1 \\ \sum_{h=1}^q \theta_h^2 \cdot V_{0,1h}^2 + \sum_{h=q+1}^p \theta_h \cdot V_{0,1h}^2 \\ \vdots \\ \sum_{h=1}^q \theta_h^M \cdot V_{0,qh}^M + \sum_{h=q+1}^p \theta_h \cdot V_{0,qh}^M \\ \sum_{h=1}^q \sum_{t=1}^M \theta_h^t \cdot V_{0,q+1h}^t + \sum_{h=q+1}^p \theta_h \cdot V_{0,q+1h} \\ \vdots \\ \sum_{h=1}^q \sum_{t=1}^M \theta_h^t \cdot V_{0,ph}^t + \sum_{h=q+1}^p \theta_h \cdot V_{0,ph} \end{pmatrix}$$

Para los cálculos que siguen voy a emplear las expresiones obtenidas del producto de

$$C_0^* \cdot C_0 = |C_0| \cdot I_c.$$

Para calcular estas expresiones voy a dividir C_0^* y C_0 en dos bloques cada una:

$$C_0^* = \begin{pmatrix} I \\ II \end{pmatrix}; \quad C_0 = \begin{pmatrix} III \\ IV \end{pmatrix}$$

El bloque I va a estar formado por las mq primeras filas de C_0^* . Voy a denominar las filas de este bloque como filas r,j , $r=1,2,\dots,M$; $j=1,2,\dots,q$. Van a estar formadas por todos los elementos de la forma $W_0^{r,jk}$, $k=1,2,\dots,p$; y de la forma $B_0^{r,j/s,k}$, $s=1,2,\dots,M$; $k=1,2,\dots,q$.

El bloque *II* va a estar formado por las $p-q$ últimas filas de C_0^* . Voy a denominar las filas de este bloque como filas $.j$, $j=q+1, \dots, p$; y estarán formadas por todos los elementos de la forma $W_0^{r,jk}$, $r=1,2,\dots,M$, $k=1,2,\dots,q$; y de la forma W_0^{jk} , $k=q+1,\dots,p$.

El bloque *III* va a estar formado por las mq primeras columnas de C_0 . Voy a denominarlas columnas $s.h$, $s=1,2,\dots,M$; $h=1,2,\dots,q$. Estarán formadas, además de por ceros, por todos los elementos de la forma $V_{0,kh}^s$, $k=1,2,\dots,p$.

El bloque *IV* va a estar formado por las $p-q$ últimas columnas de C_0 . Voy a denominarlas columnas $.h$, $h=q+1,\dots,p$. Estarán formadas por elementos de la forma $V_{0,kh}^s$, $s=1,2,\dots,M$; $k=1,2,\dots,q$; y de la forma $V_{0,kh}$, $k=q+1,\dots,p$.

Teniendo en cuenta que $C_0^* \cdot C_0 = |C_0| \cdot I_c$, se deducirá que los productos de filas de *I* por columnas de *IV*, y filas de *II* por columnas de *III* serán iguales a cero, mientras que el producto de filas de *I* por columnas de *III*, y filas de *II* por columnas de *IV* serán cero excepto en la diagonal que serán iguales a $|C_0|$. Los productos que forman la diagonal serán los de la forma *fila* $r.j \times$ *columna* $r.j$, y de la forma *fila* $.j \times$ *columna* $.j$.

Las formas que toman estos productos serán:

Bloque I \times Bloque III.

$$\text{fila } r.j \times \text{columna } r.h: \sum_{k=1}^p V_{0,kh}^r \cdot W_0^{r,jk} = \begin{cases} |C_0| & \text{si } h = j \\ 0 & \text{si } h \neq j \end{cases}$$

$$\text{fila } r.j \times \text{columna } s.h: \sum_{k=1}^q V_{0,kh}^r \cdot B_0^{r,j/s,k} + \sum_{k=q+1}^p V_{0,kh}^s \cdot W_0^{r,jk} = 0.$$

Bloque I \times Bloque IV.

$$\text{fila } r.j \times \text{columna } .h: \sum_{k=1}^q V_{0,kh}^r \cdot W_0^{r,jk} + \sum_{\substack{i=1 \\ i \neq r}}^M \sum_{k=1}^q V_{0,kh}^i \cdot B_0^{r,j/i,k} + \sum_{k=q+1}^p V_{0,kh}^r \cdot W_0^{r,jk} = 0.$$

Bloque II \times Bloque III.

$$\text{fila } .j \times \text{columna } s.h: \sum_{k=1}^q V_{0,kh}^s \cdot W_0^{s,jk} + \sum_{k=q+1}^p V_{0,kh}^s \cdot W_0^{jk} = 0$$

Bloque II \times Bloque IV.

$$\text{fila } .j \times \text{columna } .h: \sum_{i=1}^M \sum_{k=1}^q V_{0,kh}^i \cdot W_0^{i,jk} + \sum_{k=q+1}^p V_{0,kh}^r \cdot W_0^{jk} = \begin{cases} |C_0| & \text{si } h = j \\ 0 & \text{si } h \neq j \end{cases}$$

Volviendo ahora a la construcción de $C_0^{-1} D_0 = \frac{C_0^* D_0}{|C_0|}$, voy a construir primero

$C_0^* D_0$, haciéndolo por separado para los dos bloques en los que dividí C_0^* .

Para el bloque I , obtengo la convergencia casi segura de los estimadores $\hat{\theta}_j^r$, $r=1, \dots, M$;

$j=1, \dots, q$.

El estimador $\hat{\theta}_j^r$ converge casi seguro al producto *fila r , columna j* de D_0 que será igual a:

$$\begin{aligned} & \sum_{k=1}^q W_0^{r,jk} \left(\sum_{h=1}^q \theta_h' V_{0,kh}^r + \sum_{h=q+1}^p \theta_h V_{0,kh}^r \right) + \sum_{\substack{i=1 \\ i \neq r}}^M \sum_{k=1}^q B_0^{r,j/i,k} \left(\sum_{h=1}^q \theta_h' V_{0,kh}^r + \sum_{h=q+1}^p \theta_h V_{0,kh}^r \right) + \\ & + \sum_{k=q+1}^p W_0^{r,jk} \left(\sum_{h=1}^q \theta_h' V_{0,kh}^r + \sum_{h=q+1}^p \theta_h V_{0,kh}^r \right) = \\ & = \sum_{h=1}^q \theta_h' \left(\sum_{k=1}^q W_0^{r,jk} V_{0,kh}^r + \sum_{k=q+1}^p W_0^{r,jk} V_{0,kh}^r \right) + \sum_{\substack{i=1 \\ i \neq r}}^M \sum_{h=1}^q \theta_h' \left(\sum_{k=1}^q B_0^{r,j/i,k} V_{0,kh}^r + \sum_{k=q+1}^p W_0^{r,jk} V_{0,kh}^r \right) + \\ & + \sum_{h=q+1}^p \theta_h \left(\sum_{k=1}^q W_0^{r,jk} V_{0,kh}^r + \sum_{\substack{i=1 \\ i \neq r}}^M \sum_{k=1}^q B_0^{r,j/i,k} V_{0,kh}^r + \sum_{k=q+1}^p W_0^{r,jk} V_{0,kh}^r \right) = \\ & = \sum_{h=1}^q \theta_h^{(A)} + \sum_{\substack{i=1 \\ i \neq r}}^M \sum_{h=1}^q \theta_h^{(B)} + \sum_{h=q+1}^p \theta_h^{(C)} = (1) \end{aligned}$$

$$(A) = \text{fila } r.j \times \text{columna } r.h = \begin{cases} |C_0| & \text{se } h = j \\ 0 & \text{se } h \neq j \end{cases}$$

$$(B) = \text{fila } r.j \times \text{columna } t.h = 0.$$

$$(C) = \text{fila } r.j \times \text{columna } .h = 0.$$

$$\text{Entonces (1)} = \theta_h |C_0| + \sum_{\substack{h=1 \\ h \neq j}}^q \theta_h 0 + \sum_{\substack{t=1 \\ t \neq r}}^M \sum_{h=1}^q \theta_h 0 + \sum_{h=q+1}^p \theta_h 0 = \theta_h |C_0|.$$

$$\text{Con esto obtenemos que un estimador } \hat{\theta}_j \xrightarrow{\text{c.s.}} \frac{\theta_j |C_0|}{|C_0|} = \theta_j.$$

Para el bloque II, obtengo la convergencia casi segura de los estimadores $\hat{\theta}_j^1, j=q+1, \dots, p$.

El estimador $\hat{\theta}_j$ converge casi seguro al producto $\text{fila } .j \times D_0$ que será igual a:

$$\begin{aligned} & \sum_{k=1}^q \sum_{t=1}^M W_0^{r,jk} \left(\sum_{h=1}^q \theta_h V_{0,kh}^t + \sum_{h=q+1}^p \theta_h V_{0,kh}^r \right) + \sum_{k=q+1}^p W_0^{jk} \left(\sum_{h=1}^q \sum_{t=1}^M \theta_h V_{0,kh}^t + \sum_{h=q+1}^p \theta_h V_{0,kh} \right) = \\ & = \sum_{t=1}^M \sum_{h=1}^q \theta_h \left(\sum_{k=1}^q W_0^{r,jk} V_{0,kh}^r + \sum_{k=q+1}^p W_0^{jk} V_{0,kh}^r \right) + \sum_{h=q+1}^p \theta_h \left(\sum_{t=1}^M \sum_{k=1}^q W_0^{t,jk} V_{0,kh}^t + \sum_{k=q+1}^p W_0^{jk} V_{0,kh} \right) = \\ & = \sum_{t=1}^M \sum_{h=1}^q \theta_h (D) + \sum_{h=q+1}^p \theta_h (E) = (2) \end{aligned}$$

$$(D) = \text{fila } j \times \text{columna } t.h = 0.$$

$$(E) = \text{fila } j \times \text{columna } .h = \begin{cases} |C_0| & \text{se } h = j \\ 0 & \text{se } h \neq j \end{cases}$$

$$\text{Entonces (2)} = \sum_{i=1}^M \sum_{h=1}^q \theta_h^r \cdot 0 + \sum_{\substack{h=q+1 \\ h \neq j}}^p \theta_h \cdot 0 + \theta_j |C_0| = \theta_j |C_0|$$

$$\text{Con esto obtenemos que un estimador } \hat{\theta}_j \xrightarrow{c.s.} = \frac{\theta_j |C_0|}{|C_0|} = \theta_j.$$

Así demostraríamos el **apartado (1)** del teorema, o sea, demostraríamos que:

$$\hat{\Theta} \xrightarrow{c.s.} \Theta.$$

(apartado 2)

Teniendo en cuenta que $\hat{\Theta} = C^{-1} D$, podemos expresar cada uno de los estimadores

$\hat{\theta}_i^*$ como una función de los elementos de C y de D.

$$\hat{\theta}_i^* = \frac{\sum_{k=1}^c c^{ik} d_k}{|C|} = g_i^*(\dots, c_{ij}, \dots, d_k, \dots)$$

Además, eliminando los ceros de la matriz C, $\hat{\theta}_i^*$ queda como función de los

elementos de la forma $\bar{V}_{ij}^r, \bar{V}_{ij}, \bar{T}_k^s, \bar{T}_k$, entonces la demostración del **lema (2.3)**

nos permite la aplicación del *método de Bishop* al vector:

$$(\dots, \bar{V}_{ij}^r - V_{0,ij}^r, \dots, \bar{V}_{ij} - V_{0,ij}, \dots, \bar{T}_k^s - T_{0,k}^s, \dots, \bar{T}_k - T_{0,k}, \dots)'$$

lo que demuestra que $\sqrt{n}[\hat{\Theta} - \Theta] \xrightarrow{d} N_c(0_c, B' A B)$, siendo B la matriz

$$\begin{pmatrix} \frac{\partial g_1^*}{\partial \bar{V}_{ij}^s} & \frac{\partial g_h^*}{\partial \bar{V}_{ij}^s} & \frac{\partial g_c^*}{\partial \bar{V}_{ij}^s} \\ \frac{\partial g_1^*}{\partial \bar{V}_{ij}} & \frac{\partial g_h^*}{\partial \bar{V}_{ij}} & \frac{\partial g_c^*}{\partial \bar{V}_{ij}} \\ \frac{\partial g_1^*}{\partial \bar{T}_i^s} & \frac{\partial g_h^*}{\partial \bar{T}_i^s} & \frac{\partial g_c^*}{\partial \bar{T}_i^s} \\ \frac{\partial g_1^*}{\partial \bar{T}_i} & \frac{\partial g_h^*}{\partial \bar{T}_i} & \frac{\partial g_c^*}{\partial \bar{T}_i} \end{pmatrix}$$

Aplicada en el punto $(\dots, V_{0,ij}^r, \dots, V_{0,ij}, \dots, T_{0,k}^s, \dots, T_{0,k}, \dots)$

Teniendo en cuenta que $g_i^* = \frac{\sum_{k=1}^c c^{ik} d_k}{|C|}$, las derivadas serán:

$$\frac{\partial g_i^*}{\partial d_j} = \frac{c^{ji}}{|C|}; \quad \frac{\partial g_i^*}{\partial x_n} = -\frac{\left(\sum_{k=1}^c c^{ki} d_k\right) \cdot c^{in}}{|C|^2}$$

$$j \neq i \quad \frac{\partial g_i^*}{\partial x_j} = -\frac{\left(\sum_{k=1}^c c^{ki} d_k\right) \cdot |C| - \left(\sum_{k=1}^c c^{ki} d_k\right) \cdot c^{ij}}{|C|^2}$$

$c^{ki,lj}$ es el adjunto del menor $\begin{pmatrix} c_{ki} & c_{kj} \\ c_{li} & c_{lj} \end{pmatrix}$

Ahora hay que sustituir los elementos c_{ij}^0 y d_k^0 en las expresiones de las derivadas,

lo que se hace teniendo en cuenta que $d_k^0 = \sum_{x=1}^c \theta_x^* c_{kx}^0$, y que el adjunto de un

elemento c_{kl}^0 , será similar al adjunto de c_{kl} pero cambiando los elementos c_{ij} por

elementos c_{ij}^0 .

$$\left(\frac{\partial g_i^*}{\partial d_j}\right)_{d_k^0} = \frac{c^{0,ji}}{|C_0|} = c_{ji}^* \quad \left(\frac{\partial g_i^*}{\partial x_j}\right)_{c_{ij}^0} = -c_{li}^* \theta_j^*$$

c_{li}^* es el elemento (l, i) de la matriz $(C_0)^{-1}$.

La obtención de la primera expresión es inmediata, la segunda sale de la siguiente operación:

$$\begin{aligned} \left(\frac{\partial g_i^*}{\partial x_i} \right)_{c_0^0} &= - \frac{\left(\sum_{k=1}^c c^{0,ki} d_k^0 \right) \cdot c^{0,li}}{|C_0|^2} = - \frac{\left(\sum_{k=1}^c c^{0,ki} \left(\sum_{x=1}^c \theta_x^* c_{kx}^0 \right) \right) \cdot c^{0,li}}{|C_0|^2} = \\ &= - \frac{\left(\sum_{x=1}^c \theta_x^* \left(\sum_{k=1}^c c^{0,ki} c_{kx}^0 \right) \right) \cdot c^{0,li}}{|C_0|^2} = - \frac{(\theta_i^* |C_0|) \cdot c^{0,li}}{|C_0|^2} = - \frac{\theta_i^* \cdot c^{0,li}}{|C_0|} = - c_{li}^* \theta_i^* \end{aligned}$$

Ya que $\sum_{k=1}^c c^{0,ki} c_{kx}^0 = \begin{cases} |C_0| & \text{si } i = x \\ 0 & \text{si } i \neq x \end{cases}$

$$j \neq i \quad \left(\frac{\partial g_i^*}{\partial x_j} \right)_{c_0^0} = - \frac{\left(\sum_{\substack{k=1 \\ k \neq l}}^c c^{0,ki,lj} d_k^0 \right) \cdot |C_0| - \left(\sum_{k=1}^c c^{0,ki} d_k^0 \right) \cdot c^{0,lj}}{|C_0|^2} = - \frac{\left(\sum_{k=1}^c c^{0,ki} \left(\sum_{x=1}^c \theta_x^* c_{kx}^0 \right) \right) \cdot c^{0,lj}}{|C_0|^2} =$$

$$\sum_{\substack{k=1 \\ k \neq l}}^c c^{0,ki,lj} d_k^0 = \sum_{\substack{k=1 \\ k \neq l}}^c c^{0,ki,lj} \left(\sum_{x=1}^c \theta_x^* c_{kx}^0 \right) = \sum_{x=1}^c \theta_x^* \left(\sum_{\substack{k=1 \\ k \neq l}}^c c^{0,ki,lj} c_{kx}^0 \right)$$

Teniendo en cuenta que $\sum_{k=1}^c c^{0,ki,lj} c_{kx}^0 = \begin{cases} c^{0,lj} & \text{si } i = x \\ c^{0,li} & \text{si } j = x \\ 0 & \text{si } i, j \neq x \end{cases}$

$$\sum_{\substack{k=1 \\ k \neq i}}^c c^{0,ki,lj} d_k^0 = \theta_i^* c^{0,lj} + \theta_j^* c^{0,li}$$

Operando como en el caso $j=i$ se tiene: $\sum_{\substack{k=1 \\ k \neq i}}^c c^{0,ki} d_k^0 = \theta_i^* |C_0|$

Entonces:

$$\left(\frac{\partial g_i^*}{\partial c_{ij}^0} \right)_{c_{ij}^0} = - \frac{(\theta_i^* c^{0,lj} - \theta_j^* c^{0,li}) \cdot |C_0| - (\theta_i^* |C_0|) \cdot c^{0,lj}}{|C_0|^2} = - \frac{\theta_j^* c^{0,li}}{|C_0|} = - c_{li}^* \theta_j^*$$

Los elementos de la matriz B , van a ser de distintas formas. Se pueden dividir las

funciones g_h^* en funciones g_i^s cuando corresponden a un parámetro θ_i^s , con $s=1, \dots, m$;

$i=1, \dots, q$; o funciones g_i cuando corresponden a θ_i con $i=q+1, \dots, p$.

Teniendo en cuenta esas separaciones y las relaciones de los elementos \bar{V} , \bar{T} con los

elementos c_{ij} y d_j respectivamente, las formas de los elementos de la matriz B serán:

$$\left. \begin{matrix} s=1, \dots, M \\ r=1, \dots, q \\ t=1, \dots, M \\ k=1, \dots, q \end{matrix} \right\} \frac{\partial g_r^s}{\partial \bar{T}_k} = \frac{\partial g_{(s-1)q+r}^*}{\partial d_{(t-1)q+k}} = - \frac{c_{(t-1)q+k; (s-1)q+r}^*}{c_{(t-1)q+k; (s-1)q+r}} = \begin{cases} i = s & -W_{s,kr}^* \\ i \neq s & -B_{t,k/s,r}^* \end{cases}$$

$$\left. \begin{matrix} s=1, \dots, M \\ r=1, \dots, q \\ k=q+1, \dots, p \end{matrix} \right\} \frac{\partial g_r^s}{\partial \bar{T}_k} = \frac{\partial g_{(s-1)q+r}^*}{\partial d_{(M-1)q+k}} = - \frac{c_{(M-1)q+k; (s-1)q+r}^*}{c_{(M-1)q+k; (s-1)q+r}} = W_{s,kr}^*$$

$$\left. \begin{array}{l} r=q+1, \dots, p \\ i=1, \dots, M \\ k=1, \dots, q \end{array} \right\} \frac{\partial g_r}{\partial \bar{T}_k} = \frac{\partial g_{(M-1)q+r}^*}{\partial d_{(i-1)q+k}} = -C_{(i-1)q+k; (M-1)q+r}^* = W_{i,kr}^*$$

$$\left. \begin{array}{l} r=q+1, \dots, p \\ k=1, \dots, q \end{array} \right\} \frac{\partial g_r}{\partial \bar{T}_k} = \frac{\partial g_{(M-1)q+r}^*}{\partial d_{(m-1)q+k}} = -C_{(M-1)q+k; (M-1)q+r}^* = W_{kr}^*$$

$$\left. \begin{array}{l} s=1, \dots, M \\ r=1, \dots, q \\ i=1, \dots, M \\ j=1, \dots, q \\ l=1, \dots, q \end{array} \right\} \frac{\partial g_r^s}{\partial \bar{V}_{jl}^i} = \frac{\partial g_{(s-1)q+r}^*}{\partial c_{(i-1)q+j; (i-1)q+l}} = -C_{(i-1)q+j; (s-1)q+r}^* \theta_{(i-1)q+l}^* = \begin{cases} i=s & -W_{s,jr}^* \theta_l^s \\ i \neq s & -B_{i,j/s,r}^* \theta_l^i \end{cases}$$

$$\left. \begin{array}{l} s=1, \dots, M \\ r=1, \dots, q \\ i=1, \dots, M \\ j=1, \dots, q \\ l=q+1, \dots, q \end{array} \right\} \frac{\partial g_r^s}{\partial \bar{V}_{jl}^i} = \frac{\partial g_{(s-1)q+r}^*}{\partial c_{(M-1)q+j; (i-1)q+l}} = -C_{(M-1)q+j; (s-1)q+r}^* \theta_{(i-1)q+l}^* = -W_{s,jr}^* \theta_l^i$$

$$\left. \begin{array}{l} s=1, \dots, M \\ r=1, \dots, q \\ i=1, \dots, M \\ j=q+1, \dots, q \\ l=1, \dots, q \end{array} \right\} \frac{\partial g_r^s}{\partial \bar{V}_{jl}^i} = \frac{\partial g_{(s-1)q+r}^*}{\partial c_{(i-1)q+j; (M-1)q+l}} = -C_{(i-1)q+j; (s-1)q+r}^* \theta_{(M-1)q+l}^* = \begin{cases} i=s & -W_{s,jr}^* \theta_l^i \\ i \neq s & -B_{i,j/s,r}^* \theta_l^i \end{cases}$$

$$\left. \begin{array}{l} s=1, \dots, M \\ r=1, \dots, q \\ l=q+1, \dots, p \\ j=q+1, \dots, p \end{array} \right\} \frac{\partial g_r^s}{\partial \bar{V}_{jl}^i} = \frac{\partial g_{(s-1)q+r}^*}{\partial c_{(M-1)q+j; (M-1)q+l}} = -C_{(M-1)q+j; (s-1)q+r}^* \theta_{(M-1)q+l}^* = -W_{s,jr}^* \theta_l^i$$

$$\left. \begin{array}{l} r=q+1, \dots, p \\ i=1, \dots, M \\ l=1, \dots, q \\ j=1, \dots, q \end{array} \right\} \frac{\partial g_r}{\partial \bar{V}_{jl}^i} = \frac{\partial g_{(M-1)q+r}^*}{\partial c_{(i-1)q+j; (i-1)q+l}} = -C_{(i-1)q+j; (M-1)q+r}^* \theta_{(i-1)q+l}^* = -W_{i,jr}^* \theta_l^i$$

$$\left. \begin{array}{l} r=q+1, \dots, p \\ i=1, \dots, M \\ l=1, \dots, q \\ j=q+1, \dots, p \end{array} \right\} \frac{\partial g_r}{\partial \bar{V}_{jl}^i} = \frac{\partial g_{(M-1)q+r}^*}{\partial c_{(M-1)q+j; (i-1)q+l}} = -C_{(M-1)q+j; (M-1)q+r}^* \theta_{(i-1)q+l}^* = -W_{jr}^* \theta_l^i$$

$$\left. \begin{array}{l} r=q+1, \dots, p \\ i=1, \dots, M \\ l=q+1, \dots, p \\ j=1, \dots, q \end{array} \right\} \frac{\partial g_r}{\partial \bar{V}_{jl}^i} = \frac{\partial g_{(M-1)q+r}^*}{\partial c_{(i-1)q+j; (M-1)q+l}} = -C_{(i-1)q+j; (M-1)q+r}^* \theta_{(M-1)q+l}^* = -W_{i,jr}^* \theta_l^i$$

$$\left. \begin{array}{l} r=q+1, \dots, p \\ l=q+1, \dots, p \\ j=q+1, \dots, p \end{array} \right\} \frac{\partial g_r}{\partial \bar{V}_{jl}^i} = \frac{\partial g_{(M-1)q+r}^*}{\partial c_{(M-1)q+j; (M-1)q+l}} = -C_{(M-1)q+j; (M-1)q+r}^* \theta_{(M-1)q+l}^* = -W_{jr}^* \theta_l^i$$

La matriz B va a tener una columna por cada función g_h^* , y una fila por cada uno de los elementos \bar{V} o \bar{T} . Entonces teniendo en cuenta los cálculos anteriores, la matriz B quedaría como sigue:

fila	columna	g_r^s	g_r
V_{11}^1		$-B_{1,1/s,r}^* \theta_1^1$	$-W_{1,1r}^* \theta_1^1$
V_{1q}^1		$-B_{1,1/s,r}^* \theta_q^1$	$-W_{1,1r}^* \theta_q^1$
V_{1q+1}^1		$-B_{1,1/s,r}^* \theta_{q+1}^1$	$-W_{1,1r}^* \theta_{q+1}^1$
V_{1p}^1		$-B_{1,1/s,r}^* \theta_p^1$	$-W_{1,1r}^* \theta_p^1$
V_{21}^1		$-B_{1,2/s,r}^* \theta_1^1$	$-W_{1,2r}^* \theta_1^1$
V_{qp}^1		$-B_{1,q/s,r}^* \theta_p^1$	$-W_{1,qr}^* \theta_p^1$
V_{11}^s		$-W_{s,1r}^* \theta_1^s$	$-W_{s,1r}^* \theta_1^s$
V_{1p}^s		$-W_{s,1r}^* \theta_p^s$	$-W_{s,1r}^* \theta_p^s$
V_{21}^s		$-W_{s,2r}^* \theta_1^s$	$-W_{s,2r}^* \theta_1^s$
V_{qp}^s		$-W_{s,qr}^* \theta_p^s$	$-W_{s,qr}^* \theta_p^s$
V_{11}^M		$-B_{M,1/s,r}^* \theta_1^M$	$-W_{M,1r}^* \theta_1^M$
V_{1p}^M		$-B_{M,1/s,r}^* \theta_p^M$	$-W_{M,1r}^* \theta_p^M$
V_{21}^M		$-B_{M,2/s,r}^* \theta_1^M$	$-W_{M,2r}^* \theta_1^M$
V_{qp}^M		$-B_{M,q/s,r}^* \theta_p^M$	$-W_{M,qr}^* \theta_p^M$
V_{q+11}^1		$-W_{s,q+1r}^* \theta_1^1$	$-W_{q+1r}^* \theta_1^1$
V_{q+1q}^M		$-W_{s,q+1r}^* \theta_q^M$	$-W_{q+1r}^* \theta_q^M$
V_{q+1q+1}^1		$-W_{s,q+1r}^* \theta_{q+1}^1$	$-W_{q+1r}^* \theta_{q+1}^1$
V_{q+1p}^1		$-W_{s,q+1r}^* \theta_p^1$	$-W_{q+1r}^* \theta_p^1$
V_{p1}^1		$-W_{s,pr}^* \theta_1^1$	$-W_{pr}^* \theta_1^1$
V_{pq}^M		$-W_{s,pr}^* \theta_q^M$	$-W_{pr}^* \theta_q^M$
V_{pq+1}^1		$-W_{s,pr}^* \theta_{q+1}^1$	$-W_{pr}^* \theta_{q+1}^1$
V_{pp}^1		$-W_{s,pr}^* \theta_p^1$	$-W_{pr}^* \theta_p^1$
T_1^1		$B_{1,1/s,r}^*$	$W_{1,1r}^*$
T_q^1		$B_{1,q/s,r}^*$	$W_{1,qr}^*$
T_1^s		$W_{s,1r}^*$	$W_{s,1r}^*$
T_q^s		$W_{s,qr}^*$	$W_{s,qr}^*$
T_q^M		$B_{M,q/s,r}^*$	$W_{M,qr}^*$
T_{q+1}^1		$W_{s,q+1r}^*$	W_{q+1r}^*
T_p		$W_{s,pr}^*$	W_{pr}^*

Los $W_{i,jr}^*$, y los W_{jr}^* son elementos de la matriz C^{-1} , correspondientes a los elementos \bar{V}_{ij}^r , \bar{V}_{ij} de la matriz C . Los $B_{i,jl}^*$ son elementos de C^{-1} correspondientes a ceros de C , situados en la misma fila que los elementos de la forma \bar{V}_{jx}^r , y en la misma columna que los elementos de la forma \bar{V}_{yr}^s , con $x=1, 2, \dots, p$; e $y=1, 2, \dots, p$.

Voy a realizar ahora el producto $B^t A$. Para hacerlo tendré en cuenta que las columnas de A van a corresponder con los elementos \bar{V} y \bar{T} , por lo que las puedo dividir en cuatro tipos diferentes, mientras que las filas de B^t son las columnas de B , y por lo tanto de dos tipos, g_i^s y g_i . En total los elementos de $B^t A$ se obtienen de ocho tipos de productos de columnas.

$$\text{Columna } \bar{V}_{jl}^i, \text{ de } A: \left(\begin{array}{c|c|c} j=1, \dots, q & j=q+1, \dots, p & j=1, \dots, q \\ \hline l=1, \dots, q & l=1, \dots, q & l=q+1, \dots, p \end{array} \right) (i = 1, \dots, M)$$

Tiene cero los elementos correspondientes a las posiciones de los \bar{T}_h^t y de \bar{V}_{hk}^t , para $t \neq i$, siendo los demás elementos de las formas siguientes:

$$\text{Elemento } \bar{V}_{hk}^i: \text{Cov}(I^1(x)x^h x^k, x^j x^i) = C_{hk,jl}^i \cdot \left(\begin{array}{c|c|c} h=1, \dots, q & h=q+1, \dots, p & h=1, \dots, q \\ \hline k=1, \dots, q & k=1, \dots, q & k=q+1, \dots, p \end{array} \right)$$

(A partir de ahora empleare esta notación para referirme a las covarianzas)

$$\text{Elemento } \bar{V}_{hk} : C_{hk,jl}^i \quad \begin{pmatrix} h=q+1, \dots, p \\ k=q+1, \dots, p \end{pmatrix}$$

$$\text{Elemento } \bar{T}_h^i : \sum_{k=1}^q \theta_k^i C_{hk,jl}^i + \sum_{k=q+1}^p \theta_k^i C_{hk,jl}^i \quad (h = 1, \dots, q)$$

$$\text{Elemento } \bar{T}_h : \sum_{k=1}^q \theta_k^i C_{hk,jl}^i + \sum_{k=q+1}^p \theta_k^i C_{hk,jl}^i \quad (h = q+1, \dots, p)$$

El producto por una columna de una función g_r^s es el siguiente:

Cuando $s \neq i$:

$$\begin{aligned} & - \left[\sum_{h=1}^q \left(\sum_{k=1}^q C_{hk,jl}^i B_{i,h/s,r}^* \theta_k^i + \sum_{k=q+1}^p C_{hk,jl}^i B_{i,h/s,r}^* \theta_k^i \right) \right] + \left[\sum_{h=q+1}^p \left(\sum_{k=1}^q C_{hk,jl}^i W_{s,hr}^* \theta_k^i + \sum_{k=q+1}^p C_{hk,jl}^i W_{s,hr}^* \theta_k^i \right) \right] + \\ & + \left[\sum_{h=1}^q \left(\sum_{k=1}^q \theta_k^i C_{hk,jl}^i B_{i,h/s,r}^* + \sum_{k=q+1}^p \theta_k^i C_{hk,jl}^i B_{i,h/s,r}^* \right) \right] + \left[\sum_{h=q+1}^p \left(\sum_{k=1}^q \theta_k^i C_{hk,jl}^i W_{s,hr}^* + \sum_{k=q+1}^p \theta_k^i C_{hk,jl}^i W_{s,hr}^* \right) \right] = 0 \end{aligned}$$

Cuando $s=i$:

$$\begin{aligned} & - \left[\sum_{h=1}^q \left(\sum_{k=1}^q C_{hk,jl}^i W_{s,hr}^* \theta_k^i + \sum_{k=q+1}^p C_{hk,jl}^i W_{s,hr}^* \theta_k^i \right) \right] + \left[\sum_{h=q+1}^p \left(\sum_{k=1}^q C_{hk,jl}^i W_{s,hr}^* \theta_k^i + \sum_{k=q+1}^p C_{hk,jl}^i W_{s,hr}^* \theta_k^i \right) \right] + \\ & - \left[\sum_{h=1}^q \left(\sum_{k=1}^q \theta_k^i C_{hk,jl}^i W_{s,hr}^* + \sum_{k=q+1}^p \theta_k^i C_{hk,jl}^i W_{s,hr}^* \right) \right] + \left[\sum_{h=q+1}^p \left(\sum_{k=1}^q \theta_k^i C_{hk,jl}^i W_{s,hr}^* + \sum_{k=q+1}^p \theta_k^i C_{hk,jl}^i W_{s,hr}^* \right) \right] + \end{aligned}$$

El producto por una columna g_r es el siguiente:

$$-\left[\sum_{h=1}^q\left(\sum_{k=1}^q c_{hk,jl}^i W_{s,hr}^* \theta_k + \sum_{k=q+1}^p c_{hk,jl}^i W_{s,hr}^* \theta_k\right)\right] + \left[\sum_{h=q+1}^p\left(\sum_{k=1}^q c_{hk,jl}^i W_{s,hr}^* \theta_k + \sum_{k=q+1}^p c_{hk,jl}^i W_{s,hr}^* \theta_k\right)\right] +$$

$$+\left[\sum_{h=1}^q\left(\sum_{k=1}^q \theta_k^i c_{hk,jl}^i W_{s,hr}^* + \sum_{k=q+1}^p \theta_k^i c_{hk,jl}^i W_{s,hr}^*\right)\right] + \left[\sum_{h=q+1}^p\left(\sum_{k=1}^q \theta_k^i c_{hk,jl}^i W_{s,hr}^* + \sum_{k=q+1}^p \theta_k^i c_{hk,jl}^i W_{s,hr}^*\right)\right] = 0$$

Columna \bar{V}_{jl} , de \mathcal{A} $\begin{pmatrix} j=q+1, \dots, p \\ l=q+1, \dots, p \end{pmatrix}$

Elemento $\bar{V}_{hk}^t = C_{hk,jl}^t \begin{pmatrix} h=1, \dots, q & h=q+1, \dots, p & h=1, \dots, q \\ k=1, \dots, q & k=1, \dots, q & k=q+1, \dots, p \end{pmatrix} \quad (t = 1, \dots, M)$

Elemento $\bar{V}_{hk} = C_{hk,jl} \begin{pmatrix} h=q+1, \dots, p \\ k=q+1, \dots, p \end{pmatrix}$

Elemento $\bar{T}_h^t = \sum_{k=1}^q \theta_k^t c_{hk,jl}^t + \sum_{k=q+1}^p \theta_k^t c_{hk,jl}^t \quad (h = 1, \dots, q) \quad (t = 1, \dots, t)$

Elemento $\bar{T}_h = \sum_{k=1}^q \sum_{d=1}^M \theta_k^d c_{hk,jl}^d + \sum_{k=q+1}^p \theta_k^d c_{hk,jl}^d \quad (h = q+1, \dots, p)$

El producto por una columna de una función g_r^s es el siguiente:

$$-\sum_{\substack{t=1 \\ t \neq s}}^M \sum_{h=1}^q \left(\sum_{k=1}^q c_{hk,jl}^t B_{t,h/s,r}^* \theta_k^t + \sum_{k=q+1}^p c_{hk,jl}^t B_{t,h/s,r}^* \theta_k^t \right) -$$

$$\begin{aligned}
 & - \sum_{h=1}^q \left(\sum_{k=1}^q c_{hk,jl}^s W_{s,hr}^* \theta_k^s + \sum_{k=q+1}^p c_{hk,jl}^s B_{s,hr}^* \theta_k \right) - \\
 & - \sum_{h=q+1}^p \left(\sum_{d=1}^M \sum_{k=1}^q c_{hk,jl}^d W_{d,hr}^* \theta_k^d + \sum_{k=q+1}^p c_{hk,jl} W_{s,hr}^* \theta_k \right) + \\
 & + \sum_{\substack{t=1 \\ t \neq s}}^M \sum_{h=1}^q \left(\sum_{k=1}^q c_{hk,jl}^t B_{t,h/s,r}^* \theta_k^t + \sum_{k=q+1}^p c_{hk,jl}^t B_{t,h/s,r}^* \theta_k \right) + \\
 & + \sum_{h=1}^q \left(\sum_{k=1}^q c_{hk,jl}^s W_{s,hr}^* \theta_k^s + \sum_{k=q+1}^p c_{hk,jl}^s B_{s,hr}^* \theta_k \right) + \\
 & + \sum_{h=q+1}^p \left(\sum_{d=1}^M \sum_{k=1}^q c_{hk,jl}^d W_{d,hr}^* \theta_k^d + \sum_{k=q+1}^p c_{hk,jl} W_{s,hr}^* \theta_k \right) = 0
 \end{aligned}$$

El producto por una columna g_r es el siguiente:

$$\begin{aligned}
 & - \sum_{\substack{t=1 \\ t \neq s}}^M \sum_{h=1}^q \left(\sum_{k=1}^q c_{hk,jl}^t W_{t,hr}^* \theta_k^t + \sum_{k=q+1}^p c_{hk,jl}^t W_{t,hr}^* \theta_k \right) - \sum_{h=q+1}^p \left(\sum_{t=1}^M \sum_{k=1}^q c_{hk,jl}^t W_{t,hr}^* \theta_k^t + \sum_{k=q+1}^p c_{hk,jl} W_{hr}^* \theta_k \right) + \\
 & + \sum_{\substack{t=1 \\ t \neq s}}^M \sum_{h=1}^q \left(\sum_{k=1}^q c_{hk,jl}^t W_{t,hr}^* \theta_k^t + \sum_{k=q+1}^p c_{hk,jl}^t W_{t,hr}^* \theta_k \right) + \sum_{h=q+1}^p \left(\sum_{t=1}^M \sum_{k=1}^q c_{hk,jl}^t W_{t,hr}^* \theta_k^t + \sum_{k=q+1}^p c_{hk,jl} W_{hr}^* \theta_k \right) = 0
 \end{aligned}$$

Columna \bar{T}_j^t , de \mathcal{A} : $(j = 1, \dots, q) \quad (i = 1, \dots, M)$

Tiene cero los elementos correspondientes a las posiciones de los \bar{T}_h^t y de los \bar{V}_{hk}^t ,

para $t \neq i$, siendo los demás elementos de las formas siguientes:

$$\text{Elemento } \bar{V}_{hk} : \sum_{v=1}^q \theta_v^i c_{hk,jv}^i + \sum_{v=q+1}^p \theta_v^i c_{hk,jv}^i \begin{pmatrix} h=1,\dots,q & h=q+1,\dots,p & h=1,\dots,q \\ k=1,\dots,q & k=1,\dots,q & k=q+1,\dots,p \end{pmatrix}$$

$$\text{Elemento } \bar{V}_{hk} : \sum_{v=1}^q \theta_v^i c_{hk,jv}^i + \sum_{v=q+1}^p \theta_v^i c_{hk,jv}^i \begin{pmatrix} h=q+1,\dots,p \\ k=q+1,\dots,p \end{pmatrix}$$

$$\text{Elemento } \bar{T}_h^i : \sum_{v=1}^q \sum_{u=1}^q \theta_u^i \theta_v^i c_{hv,ju}^i + \sum_{v=1}^q \sum_{u=q+1}^p \theta_v^i \theta_u^i c_{hv,ju}^i +$$

$$\sum_{v=q+1}^p \sum_{u=1}^q \theta_u^i \theta_v^i c_{hv,ju}^i + \sum_{v=q+1}^p \sum_{u=q+1}^p \theta_v^i \theta_u^i c_{hv,ju}^i + \sigma^2 E[I^i(x) x^j x^h] \quad (h = 1, \dots, q)$$

$$\text{Elemento } \bar{T}_h^i : \sum_{v=1}^q \sum_{u=1}^q \theta_u^i \theta_v^i c_{hv,ju}^i + \sum_{v=1}^q \sum_{u=q+1}^p \theta_v^i \theta_u^i c_{hv,ju}^i +$$

$$\sum_{v=q+1}^p \sum_{u=1}^q \theta_u^i \theta_v^i c_{hv,ju}^i + \sum_{v=q+1}^p \sum_{u=q+1}^p \theta_v^i \theta_u^i c_{hv,ju}^i + \sigma^2 E[I^i(x) x^j x^h] \quad (h = q+1, \dots, p)$$

El producto por una columna de una función g_s^i es el siguiente:

Cuando $s \neq i$:

$$\begin{aligned} & - \left[\sum_{h=1}^q \left[\sum_{k=1}^q \left(\sum_{v=1}^q B_{i,h/s,r}^* \theta_k^i \theta_v^i c_{hk,jv}^i + \sum_{v=q+1}^p B_{i,h/s,r}^* \theta_k^i \theta_v^i c_{hk,jv}^i \right) + \right. \right. \\ & \quad \left. \left. + \sum_{k=q+1}^p \left(\sum_{v=1}^q B_{i,h/s,r}^* \theta_k^i \theta_v^i c_{hk,jv}^i + \sum_{v=q+1}^p B_{i,h/s,r}^* \theta_k^i \theta_v^i c_{hk,jv}^i \right) \right] + \right. \\ & \quad \left. + \sum_{h=q+1}^p \left[\sum_{k=1}^q \left(\sum_{v=1}^q W_{s,hr}^* \theta_k^i \theta_v^i c_{hk,jv}^i + \sum_{v=q+1}^p W_{s,hr}^* \theta_k^i \theta_v^i c_{hk,jv}^i \right) + \right. \right. \end{aligned}$$

$$\begin{aligned}
 & + \sum_{k=q+1}^p \left(\sum_{v=1}^q W_{s,hr}^* \theta_k \theta_v c_{hk,jv}^i + \sum_{v=q+1}^p W_{s,hr}^* \theta_k \theta_v c_{hk,jv}^i \right) \Bigg] + \\
 & + \left[\sum_{h=1}^q \left[\sum_{k=1}^q \left(\sum_{v=1}^q B_{i,h/s,r}^* \theta_k \theta_v c_{hk,jv}^i + \sum_{v=q+1}^p B_{i,h/s,r}^* \theta_k \theta_v c_{hk,jv}^i \right) + \right. \right. \\
 & \left. \left. + \sum_{k=q+1}^p \left(\sum_{v=1}^q B_{i,h/s,r}^* \theta_k \theta_v c_{hk,jv}^i + \sum_{v=q+1}^p B_{i,h/s,r}^* \theta_k \theta_v c_{hk,jv}^i \right) \right] + \right. \\
 & \left. + \sum_{h=q+1}^p \left[\sum_{k=1}^q \left(\sum_{v=1}^q W_{s,hr}^* \theta_k \theta_v c_{hk,jv}^i + \sum_{v=q+1}^p W_{s,hr}^* \theta_k \theta_v c_{hk,jv}^i \right) + \right. \right. \\
 & \left. \left. + \sum_{k=q+1}^p \left(\sum_{v=1}^q W_{s,hr}^* \theta_k \theta_v c_{hk,jv}^i + \sum_{v=q+1}^p W_{s,hr}^* \theta_k \theta_v c_{hk,jv}^i \right) \right] \right] + \sigma^2 \sum_{h=1}^p E[I^i(x) x^j x^h] W_{s,hr}^* = \\
 & 0 + \sigma^2 \sum_{h=1}^p V_{0,hj}^i W_{s,hr}^* = 0, \text{ ya que } i \neq s.
 \end{aligned}$$

Cuando $s=i$:

$$\begin{aligned}
 & - \left[\sum_{h=1}^q \left[\sum_{k=1}^q \left(\sum_{v=1}^q W_{s,hr}^* \theta_k^s \theta_v^s c_{hk,jv}^s + \sum_{v=q+1}^p W_{s,hr}^* \theta_k^s \theta_v^s c_{hk,jv}^s \right) + \right. \right. \\
 & \left. \left. + \sum_{k=q+1}^p \left(\sum_{v=1}^q W_{s,hr}^* \theta_k^s \theta_v^s c_{hk,jv}^s + \sum_{v=q+1}^p W_{s,hr}^* \theta_k^s \theta_v^s c_{hk,jv}^s \right) \right] + \right. \\
 & \left. + \sum_{h=q+1}^p \left[\sum_{k=1}^q \left(\sum_{v=1}^q W_{s,hr}^* \theta_k^s \theta_v^s c_{hk,jv}^s + \sum_{v=q+1}^p W_{s,hr}^* \theta_k^s \theta_v^s c_{hk,jv}^s \right) + \right. \right.
 \end{aligned}$$

$$\begin{aligned}
 & + \sum_{k=q+1}^p \left(\sum_{v=1}^q W_{s,hr}^* \theta_k \theta_v^s c_{hk,jv}^s + \sum_{v=q+1}^p W_{s,hr}^* \theta_k \theta_v^s c_{hk,jv}^s \right) \Bigg] + \\
 & + \left[\sum_{h=1}^q \left[\sum_{k=1}^q \left(\sum_{v=1}^q W_{s,hr}^* \theta_k \theta_v^s c_{hk,jv}^s + \sum_{v=q+1}^p W_{s,hr}^* \theta_k \theta_v^s c_{hk,jv}^s \right) + \right. \right. \\
 & + \sum_{k=q+1}^p \left(\sum_{v=1}^q W_{s,hr}^* \theta_k \theta_v^s c_{hk,jv}^s + \sum_{v=q+1}^p W_{s,hr}^* \theta_k \theta_v^s c_{hk,jv}^s \right) \Bigg] + \\
 & + \sum_{h=q+1}^p \left[\sum_{k=1}^q \left(\sum_{v=1}^q W_{s,hr}^* \theta_k \theta_v^s c_{hk,jv}^s + \sum_{v=q+1}^p W_{s,hr}^* \theta_k \theta_v^s c_{hk,jv}^s \right) + \right. \\
 & + \sum_{k=q+1}^p \left(\sum_{v=1}^q W_{s,hr}^* \theta_k \theta_v^s c_{hk,jv}^s + \sum_{v=q+1}^p W_{s,hr}^* \theta_k \theta_v^s c_{hk,jv}^s \right) \Bigg] + \sigma^2 \sum_{h=1}^p E[I^s(x) x' x^h] W_{s,hr}^* = \\
 & 0 + \sigma^2 \sum_{h=1}^p V_{0,hj}^s W_{s,hr}^* = \begin{cases} 0, & \text{cuando } j \neq r \\ \sigma^2, & \text{cuando } j = r \end{cases}
 \end{aligned}$$

El producto por una columna g_r es el siguiente:

$$\begin{aligned}
 & - \left[\sum_{h=1}^q \left[\sum_{k=1}^q \left(\sum_{v=1}^q W_{i,hr}^* \theta_k \theta_v^i c_{hk,jv}^i + \sum_{v=q+1}^p W_{i,hr}^* \theta_k \theta_v^i c_{hk,jv}^i \right) + \right. \right. \\
 & + \sum_{k=q+1}^p \left(\sum_{v=1}^q W_{i,hr}^* \theta_k \theta_v^i c_{hk,jv}^i + \sum_{v=q+1}^p W_{i,hr}^* \theta_k \theta_v^i c_{hk,jv}^i \right) \Bigg] + \\
 & + \sum_{h=q+1}^p \left[\sum_{k=1}^q \left(\sum_{v=1}^q W_{i,hr}^* \theta_k \theta_v^i c_{hk,jv}^i + \sum_{v=q+1}^p W_{i,hr}^* \theta_k \theta_v^i c_{hk,jv}^i \right) + \right. \\
 & + \sum_{k=q+1}^p \left(\sum_{v=1}^q W_{i,hr}^* \theta_k \theta_v^i c_{hk,jv}^i + \sum_{v=q+1}^p W_{i,hr}^* \theta_k \theta_v^i c_{hk,jv}^i \right) \Bigg] +
 \end{aligned}$$

$$\begin{aligned}
 & + \sum_{k=q+1}^p \left(\sum_{v=1}^q W_{hr}^* \theta_k \theta_v^i c_{hk,jv}^i + \sum_{v=q+1}^p W_{hr}^* \theta_k \theta_v^i c_{hk,jv}^i \right) \Bigg] + \\
 & + \left[\sum_{h=1}^q \left[\sum_{k=1}^q \left(\sum_{v=1}^q W_{i,hr}^* \theta_k \theta_v^i c_{hk,jv}^i + \sum_{v=q+1}^p W_{i,hr}^* \theta_k \theta_v^i c_{hk,jv}^i \right) + \right. \right. \\
 & + \sum_{k=q+1}^p \left(\sum_{v=1}^q W_{i,hr}^* \theta_k \theta_v^i c_{hk,jv}^i + \sum_{v=q+1}^p W_{i,hr}^* \theta_k \theta_v^i c_{hk,jv}^i \right) \Bigg] + \\
 & + \sum_{h=q+1}^p \left[\sum_{k=1}^q \left(\sum_{v=1}^q W_{hr}^* \theta_k \theta_v^i c_{hk,jv}^i + \sum_{v=q+1}^p W_{hr}^* \theta_k \theta_v^i c_{hk,jv}^i \right) + \right. \\
 & + \sum_{k=q+1}^p \left(\sum_{v=1}^q W_{hr}^* \theta_k \theta_v^i c_{hk,jv}^i + \sum_{v=q+1}^p W_{hr}^* \theta_k \theta_v^i c_{hk,jv}^i \right) \Bigg] + \\
 & + \sigma^2 \left(\sum_{h=1}^q E[I^i(x) x^j x^h] W_{i,hr}^* + \sum_{h=q+1}^p E[I^i(x) x^j x^h] W_{hr}^* \right) = 0 + \sigma^2 \left(\sum_{h=1}^q V_{0,ij}^s W_{s,hr}^* + \sum_{h=q+1}^p V_{0,ij}^s W_{hr}^* \right) = \\
 & = 0 \quad \text{pues} \quad \left. \begin{array}{l} i \leq q \\ r \geq q+1 \end{array} \right\} \Rightarrow i \neq r
 \end{aligned}$$

Columna \bar{T}_j , de \mathcal{A} ($j = q+1, \dots, p$)

$$\text{Elemento } \bar{V}_{hk}^i: \sum_{v=1}^q \theta_v^i c_{hk,jv}^i + \sum_{v=q+1}^p \theta_v^i c_{hk,jv}^i \quad \left(\begin{array}{c|c|c} h=1, \dots, q & h=q+1, \dots, p & h=1, \dots, q \\ \hline k=1, \dots, q & k=1, \dots, q & k=q+1, \dots, p \end{array} \right) \quad (i=1, \dots, M)$$

$$\text{Elemento } \bar{V}_{hk}: \sum_{v=1}^q \sum_{d=1}^M \theta_v^d c_{hk,jv}^d + \sum_{v=q+1}^p \theta_v^i c_{hk,jv}^i \quad \left(\begin{array}{c} h=q+1, \dots, p \\ \hline k=q+1, \dots, p \end{array} \right)$$

$$\text{Elemento } \bar{T}_h^t: \sum_{v=1}^q \sum_{u=1}^q \theta_u^t \theta_v^t c_{hv,ju}^t + \sum_{v=1}^q \sum_{u=q+1}^p \theta_v^t \theta_u^t c_{hv,ju}^t + \\ + \sum_{v=q+1}^p \sum_{u=1}^q \theta_u^t \theta_v^t c_{hv,ju}^t + \sum_{v=q+1}^p \sum_{u=q+1}^p \theta_v^t \theta_u^t c_{hv,ju}^t + \sigma^2 E[I^t(x) x^j x^h] \quad (h = 1, \dots, q) \quad (t = 1, \dots, M)$$

$$\text{Elemento } \bar{T}_h^d: \sum_{v=1}^q \sum_{u=1}^q \sum_{d=1}^M \theta_u^d \theta_v^d c_{hv,ju}^d + \sum_{v=1}^q \sum_{u=q+1}^p \sum_{d=1}^M \theta_v^d \theta_u^d c_{hv,ju}^d + \\ + \sum_{v=q+1}^p \sum_{u=1}^q \sum_{d=1}^M \theta_u^d \theta_v^d c_{hv,ju}^d + \sum_{v=q+1}^p \sum_{u=q+1}^p \theta_v^d \theta_u^d c_{hv,ju}^d + \sigma^2 E[x^j x^h] \quad (h = q+1, \dots, p)$$

El producto por una columna de una función g_s^j es el siguiente:

$$\left[\sum_{h=1}^q \left[\sum_{\substack{t=1 \\ t \neq s}}^M \left[\sum_{k=1}^q \left(\sum_{v=1}^q B_{t,h/s,r}^* \theta_k^t \theta_v^t c_{hk,jv}^t + \sum_{v=q+1}^p B_{t,h/s,r}^* \theta_k^t \theta_v^t c_{hk,jv}^t \right) + \right. \right. \right. \\ \left. \left. + \sum_{k=q+1}^p \left(\sum_{v=1}^q B_{t,h/s,r}^* \theta_k^t \theta_v^t c_{hk,jv}^t + \sum_{v=q+1}^p B_{t,h/s,r}^* \theta_k^t \theta_v^t c_{hk,jv}^t \right) \right] + \right. \\ \left. + \sum_{k=1}^q \left(\sum_{v=1}^q W_{s,hr}^* \theta_k^s \theta_v^s c_{hk,jv}^s + \sum_{v=q+1}^p W_{s,hr}^* \theta_k^s \theta_v^s c_{hk,jv}^s \right) + \right. \\ \left. + \sum_{k=q+1}^p \left(\sum_{v=1}^q W_{s,hr}^* \theta_k^s \theta_v^s c_{hk,jv}^s + \sum_{v=q+1}^p W_{s,hr}^* \theta_k^s \theta_v^s c_{hk,jv}^s \right) \right] + \\ + \sum_{h=q+1}^p \left[\sum_{k=1}^q \left(\sum_{v=1}^q W_{hr}^* \theta_k^s \theta_v^s c_{hk,jv}^s + \sum_{v=q+1}^p W_{hr}^* \theta_k^s \theta_v^s c_{hk,jv}^s \right) + \right. \\ \left. + \sum_{k=q+1}^p \left(\sum_{v=1}^q W_{hr}^* \theta_k^s \theta_v^s c_{hk,jv}^s + \sum_{v=q+1}^p W_{hr}^* \theta_k^s \theta_v^s c_{hk,jv}^s \right) \right]$$

$$\begin{aligned}
 & + \sum_{k=q+1}^p \left(\sum_{v=1}^q W_{hr}^* \theta_k \theta_v^s c_{hk,jv}^s + \sum_{v=q+1}^p W_{hr}^* \theta_k \theta_v c_{hk,jv} \right) \Bigg] + \\
 & + \left[\sum_{h=1}^q \left[\sum_{\substack{t=1 \\ t \neq s}}^M \left[\sum_{k=1}^q \left(\sum_{v=1}^q B_{t,h/s,r}^* \theta_k \theta_v^t c_{hk,jv}^t + \sum_{v=q+1}^p B_{t,h/s,r}^* \theta_k \theta_v c_{hk,jv}^t \right) + \right. \right. \right. \\
 & + \sum_{k=q+1}^p \left(\sum_{v=1}^q B_{t,h/s,r}^* \theta_k \theta_v^t c_{hk,jv}^t + \sum_{v=q+1}^p B_{t,h/s,r}^* \theta_k \theta_v c_{hk,jv}^t \right) \Bigg] + \\
 & + \sum_{k=1}^q \left(\sum_{v=1}^q W_{s,hr}^* \theta_k \theta_v^s c_{hk,jv}^s + \sum_{v=q+1}^p W_{s,hr}^* \theta_k \theta_v c_{hk,jv}^s \right) + \\
 & + \sum_{k=q+1}^p \left(\sum_{v=1}^q W_{s,hr}^* \theta_k \theta_v^s c_{hk,jv}^s + \sum_{v=q+1}^p W_{s,hr}^* \theta_k \theta_v c_{hk,jv}^s \right) \Bigg] + \\
 & + \sum_{h=q+1}^p \left[\sum_{k=1}^q \left(\sum_{v=1}^q W_{hr}^* \theta_k \theta_v^s c_{hk,jv}^s + \sum_{v=q+1}^p W_{hr}^* \theta_k \theta_v c_{hk,jv}^s \right) + \right. \\
 & + \sum_{k=q+1}^p \left(\sum_{v=1}^q W_{hr}^* \theta_k \theta_v^s c_{hk,jv}^s + \sum_{v=q+1}^p W_{hr}^* \theta_k \theta_v c_{hk,jv}^s \right) \Bigg] + \\
 & + \sigma^2 \left(\sum_{\substack{t=1 \\ t \neq s}}^M \sum_{h=1}^q B_{t,h/s,r}^* V_{0,hj}^t + \sum_{h=1}^q W_{s,hr}^* V_{0,hj}^s + \sum_{h=q+1}^p W_{s,hr}^* V_{0,hj}^s \right) = \\
 & = 0 + \sigma^2 \left(\sum_{\substack{t=1 \\ t \neq s}}^M \sum_{h=1}^q B_{t,h/s,r}^* V_{0,hj}^t + \sum_{h=1}^q W_{s,hr}^* V_{0,hj}^s + \sum_{h=q+1}^p W_{s,hr}^* V_{0,hj}^s \right) = 0
 \end{aligned}$$

ya que $h \geq q+1$, lo que hace que los valores de V_0 estén en las $p-q$ últimas filas mientras que los valores que los multiplican corresponden siempre a filas anteriores.

El producto por una columna g_r es el siguiente:

$$\begin{aligned}
& - \left[\sum_{h=1}^q \sum_{t=1}^M \left[\sum_{k=1}^q \left(\sum_{v=1}^q W_{t,hr}^* \theta_k^t \theta_v^t c_{hk,jv}^t + \sum_{v=q+1}^p W_{t,hr}^* \theta_k^t \theta_v^t c_{hk,jv}^t \right) + \right. \right. \\
& \quad \left. \left. + \sum_{k=q+1}^p \left(\sum_{v=1}^q W_{t,hr}^* \theta_k^t \theta_v^t c_{hk,jv}^t + \sum_{v=q+1}^p W_{t,hr}^* \theta_k^t \theta_v^t c_{hk,jv}^t \right) \right] + \right. \\
& \quad \left. + \sum_{h=q+1}^p \sum_{t=1}^M \left[\sum_{k=1}^q \left(\sum_{v=1}^q W_{hr}^* \theta_k^t \theta_v^t c_{hk,jv}^t + \sum_{v=q+1}^p W_{hr}^* \theta_k^t \theta_v^t c_{hk,jv}^t \right) + \right. \right. \\
& \quad \left. \left. + \sum_{k=q+1}^p \left(\sum_{v=1}^q W_{hr}^* \theta_k^t \theta_v^t c_{hk,jv}^t + \sum_{v=q+1}^p W_{hr}^* \theta_k^t \theta_v^t c_{hk,jv}^t \right) \right] \right] + \\
& \quad + \left[\sum_{h=1}^q \sum_{t=1}^M \left[\sum_{k=1}^q \left(\sum_{v=1}^q W_{t,hr}^* \theta_k^t \theta_v^t c_{hk,jv}^t + \sum_{v=q+1}^p W_{t,hr}^* \theta_k^t \theta_v^t c_{hk,jv}^t \right) + \right. \right. \\
& \quad \left. \left. + \sum_{k=q+1}^p \left(\sum_{v=1}^q W_{t,hr}^* \theta_k^t \theta_v^t c_{hk,jv}^t + \sum_{v=q+1}^p W_{t,hr}^* \theta_k^t \theta_v^t c_{hk,jv}^t \right) \right] + \right. \\
& \quad \left. + \sum_{h=q+1}^p \sum_{t=1}^M \left[\sum_{k=1}^q \left(\sum_{v=1}^q W_{hr}^* \theta_k^t \theta_v^t c_{hk,jv}^t + \sum_{v=q+1}^p W_{hr}^* \theta_k^t \theta_v^t c_{hk,jv}^t \right) + \right. \right. \\
& \quad \left. \left. + \sum_{k=q+1}^p \left(\sum_{v=1}^q W_{hr}^* \theta_k^t \theta_v^t c_{hk,jv}^t + \sum_{v=q+1}^p W_{hr}^* \theta_k^t \theta_v^t c_{hk,jv}^t \right) \right] \right] + \\
& \quad = 0 + \sigma^2 \left(\sum_{t=1}^M \sum_{h=1}^q W_{t,hr}^* V_{0,hj}^t + \sum_{h=q+1}^p W_{hr}^* V_{0,hj} \right) = \begin{cases} 0, & \text{si } j \neq r \\ \sigma^2, & \text{si } j = r \end{cases}
\end{aligned}$$

Quedan así calculados los elementos de la matriz $B^t \mathcal{A}$, quedando esta como una matriz con tantas filas como funciones g_h^* , y tantas columnas como elementos \bar{V} y \bar{T} , pudiendo además ser dividida en dos bloques. El primero formado por las columnas correspondientes a los elementos \bar{V} y constituida sólo por ceros. El segundo corresponde a las columnas de los elementos \bar{T} , y toma la forma $\sigma^2 \cdot I_c$ (La varianza por la matriz identidad de orden c).

Si dividimos también B en dos bloques, B_V y B_T que agrupan las filas según los elementos que las determinan, tendré el siguiente producto:

$$B^t \mathcal{A} B = \begin{pmatrix} 0_v, \sigma^2 I_c \end{pmatrix} \cdot \begin{pmatrix} B_v \\ B_T \end{pmatrix} = \sigma^2 I_c B_T = \sigma^2 B_T$$

Teniendo en cuenta que B_T está formado por los elementos de la matriz inversa de

$$C_0 = E[A(X)A'(X)],$$

llegamos a demostrar que

$$\sqrt{n}(\hat{\Theta} - \Theta) \xrightarrow{d} N_c \left(0, \sigma^2 \left(E[A(X)A'(X)] \right)^{-1} \right)$$

Con lo que se completa la demostración de este teorema.

APÉNDICE III.1 Procedimientos empleados en la elaboración de los datos.

Proceso de trimestralización de las series de salarios.

- i) Como el crecimiento de este índice se puede suponer geométrico, se comenzó haciendo una regresión del logaritmo del índice ($\lg W$) con respecto del tiempo (s):

$$\lg \tilde{W} = \hat{\alpha} + \hat{\beta} s$$

- ii) Se calculan los residuos estimados para cada trimestre $t = 1, 2, 3, 4$, del año $n = 0, 1, 2, \dots, N$. (siendo 0 el año 1981):

$$\lg r_{nt} = \lg W - \hat{\alpha} - \hat{\beta}(4n + t)$$

- iii) Se calculó el promedio de los residuos para cada trimestre t :

$$\lg I_t = \frac{\sum_{n=1}^N \lg r_{nt}}{N}$$

- iv) Se procedió a centrar esos residuos, de manera que la suma de los cuatro residuos medios trimestrales estimados sea cero. El proceso sería el siguiente:

1) Calcular $k = \sum_{t=1}^4 \lg I_t$.

2) Centrar los residuos: $\lg^* I_t = \lg I_t - \frac{k}{4}$, $t = 1, 2, 3, 4$

- v) Se deshace el cambio logarítmico:

$$I_t^* = e^{\lg^* I_t}$$

El valor I_t^* es una estimación de la relación entre el incremento medio del trimestre t y la tasa de variación media de los cuatro trimestres.

Si se quiere estimar los índices trimestrales para el año n , se parte de la variación producida ese año, τ_n . El incremento que se hace corresponder a cada trimestre del año va a ser:

$$\tau_{nt}^* = \frac{\sqrt[4]{\tau_n}}{I_t^*}$$

Con esto, partiendo de un índice conocido se pueden construir los índices trimestrales para el año siguiente o anterior, teniendo sólo en cuenta que:

$$S_{n,1}^* = \tau_{n,1}^* \cdot S_{n-1,4}$$

$$S_{n,2}^* = \tau_{n,2}^* \cdot S_{n,1}^*$$

$$S_{n,3}^* = \tau_{n,3}^* \cdot S_{n,2}^*$$

$$S_{n,4}^* = \tau_{n,4}^* \cdot S_{n,3}^*$$

Una vez completada la serie de los índices se procedió a trimestralizar la serie de los salarios de la forma siguiente:

Sean:

τ_{nt} , el incremento producido a nivel nacional para los salarios en el año n y en el trimestre t .

X_{nt} , el salario en una comunidad en el año n .

- i) Se calculó el salario que tendría la comunidad en el año $n+1$, si se produce el mismo incremento que a nivel nacional:

$$X_{n+1}^* = X_n \cdot \tau_{n1} \cdot \tau_{n2} \cdot \tau_{n3} \cdot \tau_{n4}$$

- ii) Como este salario estimado debe coincidir con el que existió realmente en esa comunidad, se ajustan los incrementos trimestrales para que eso ocurra:

$$\tau_{ni}^* = \tau_{ni} \cdot \sqrt[4]{\frac{X_{n+1}}{X_{n+1}^*}}$$

Con lo cual se tiene que:

$$X_{n+1} = X_n \cdot \tau_{n1}^* \cdot \tau_{n2}^* \cdot \tau_{n3}^* \cdot \tau_{n4}^*$$

Así se tendría trimestralizada la serie de salarios entre I-80 y IV-90.

Proceso de suavización de las series de estudiantes de secundaria.

Para realizar esta suavización se definió primero una función objetivo que tuviera en cuenta la distancia entre los valores observados y los estimados, y a partir de ella se definieron unas ponderaciones para construir unos valores que se adaptasen a los datos y a su vez mantuviesen una tendencia más suavizada.

En particular para suavizar esos valores se ajustó una recta y una función exponencial en cada comunidad:

$$X_t^1 = \hat{a} + \hat{b}X_t$$

$$X_t^2 = \hat{a} e^{\hat{b}X_t}$$

A partir de esas estimaciones se construyeron unos valores

$$\tilde{X}_i^j = (1 - \alpha_i) X_i^{j*} + \alpha_i X_i^j \quad i = 1, 2, \dots$$

siendo $\alpha_i \in [0, 1]$, el valor que minimiza la función $\sum_{i=1}^T (\tilde{X}_i^j - X_i^j)^2 \quad i = 1, 2, \dots$

verificando: $\tilde{X}_i^j \geq \tilde{X}_{i-1}^j$

Los valores X_i anómalos fueron sustituidos por aquel de los correspondientes \tilde{X}_i^j que presentaba un menor valor en la función objetivo

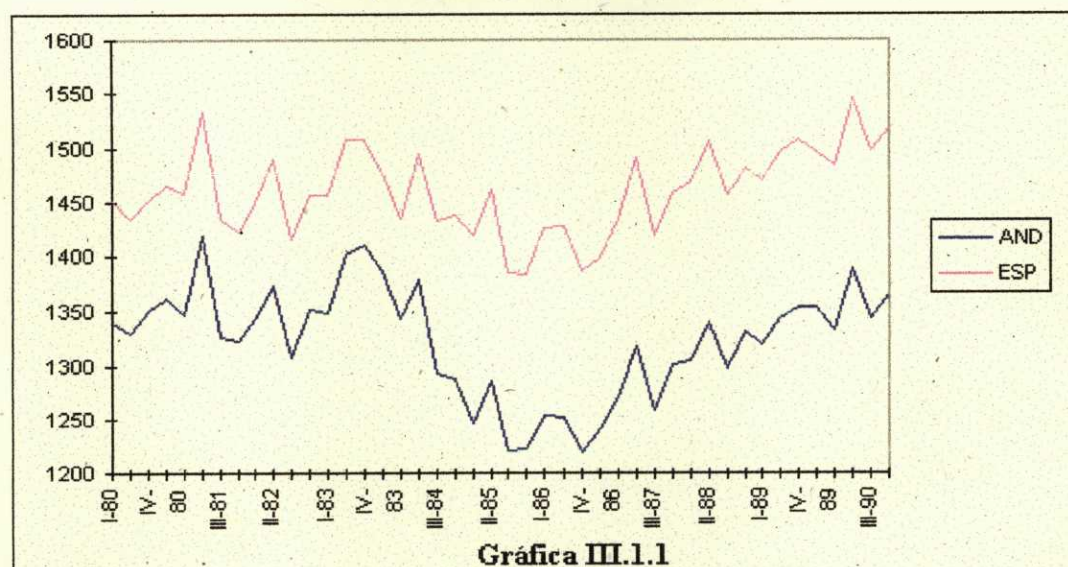
APÉNDICE III.2 Datos empleados.

En este apéndice se incluyen los datos elaborados a partir de los originales (*Apartado 3.2.2 y Apartado 3.5*). A partir de ellos se realizarán los estudios consignados en el *Capítulo 4* de este trabajo.

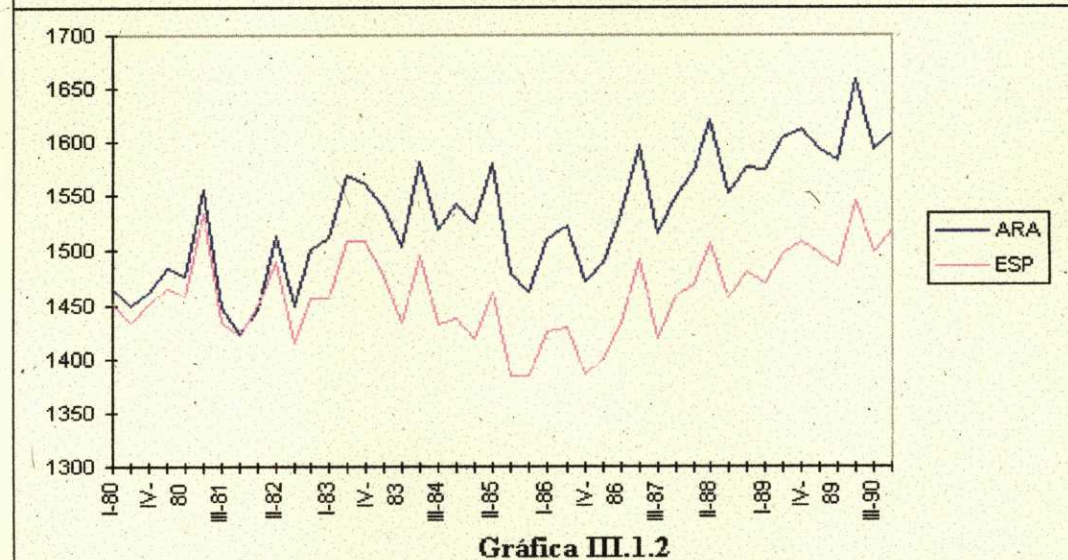
El texto de este apéndice está incluido en los ficheros APIII2.doc (formato Microsoft Word 6.0) y APIII2.wp (formato Word Perfect 5.1).

APÉNDICE III.3 Gráficas de las variables.

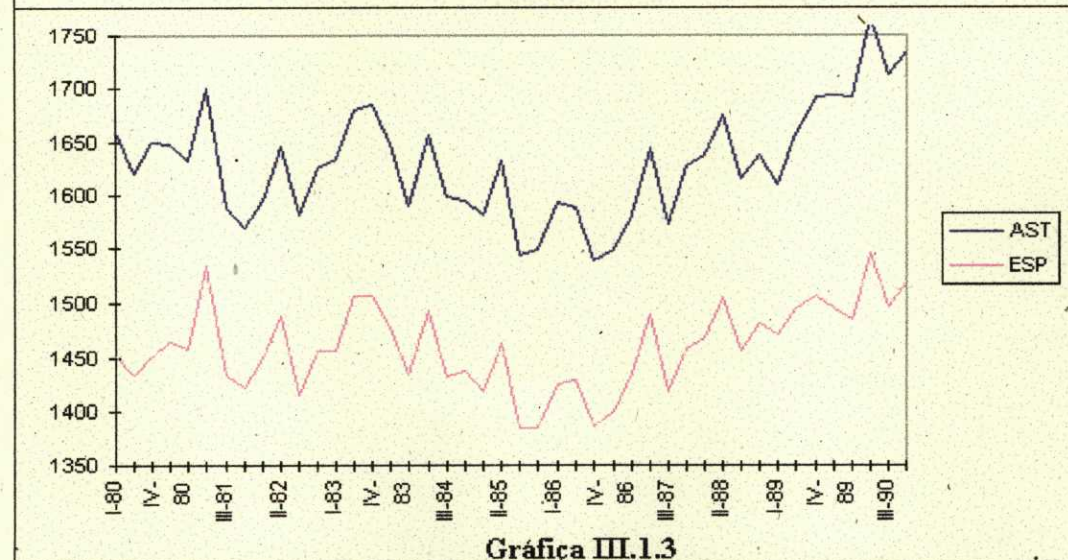
Nivel de Salarios. (Andalucía, Aragón y Asturias)



Gráfica III.1.1

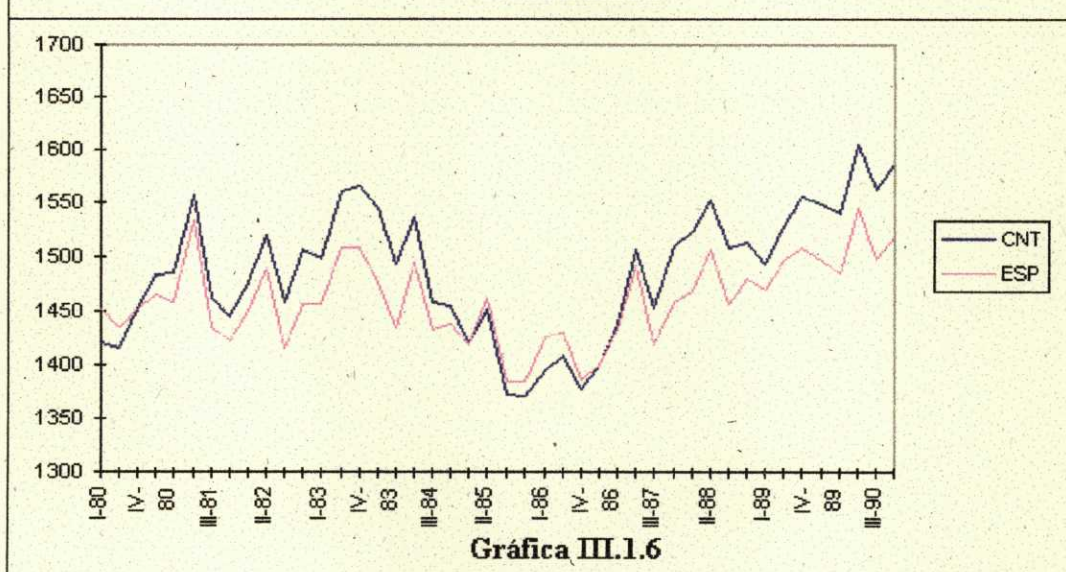
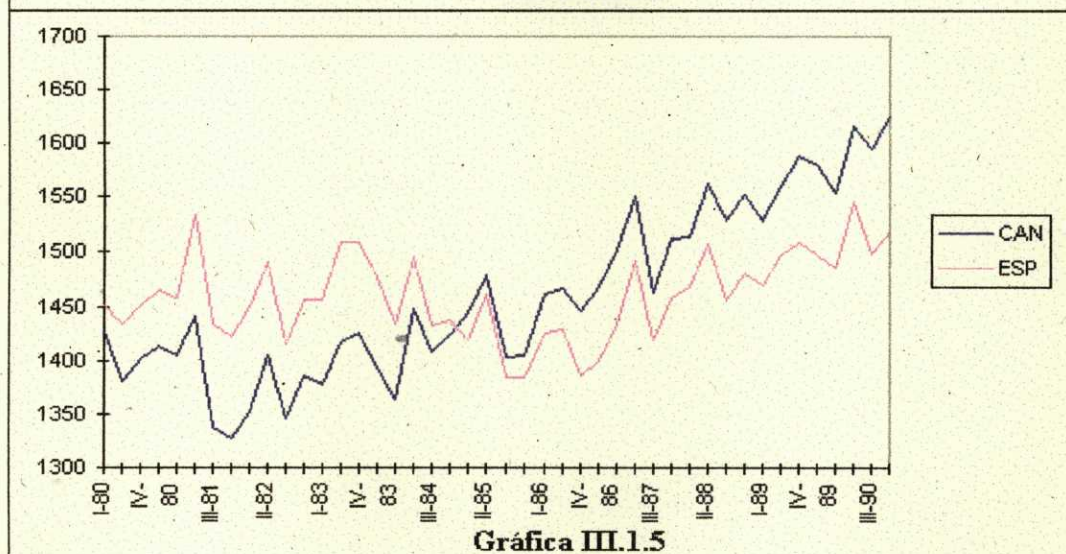
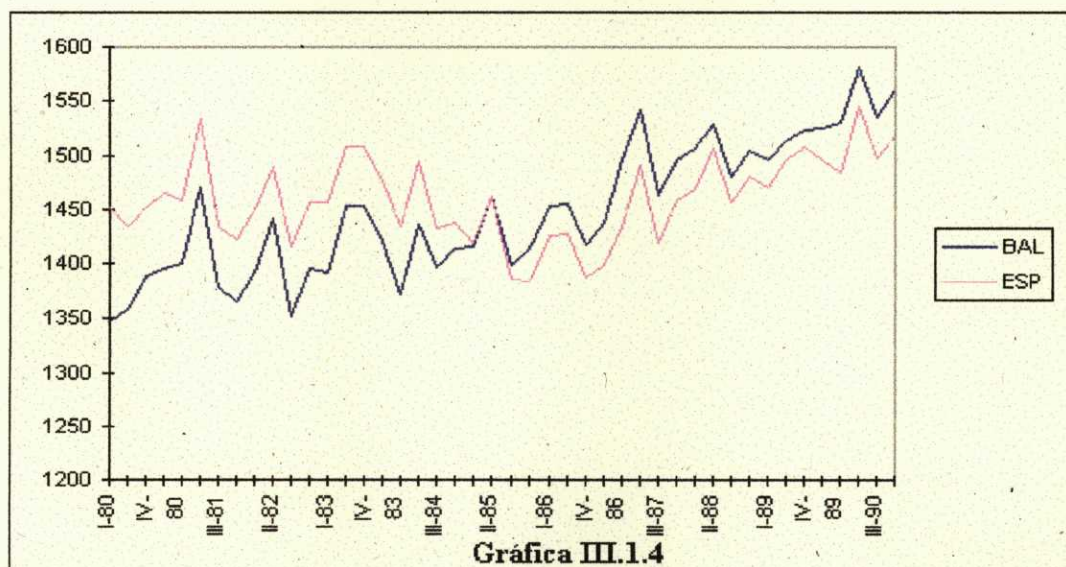


Gráfica III.1.2

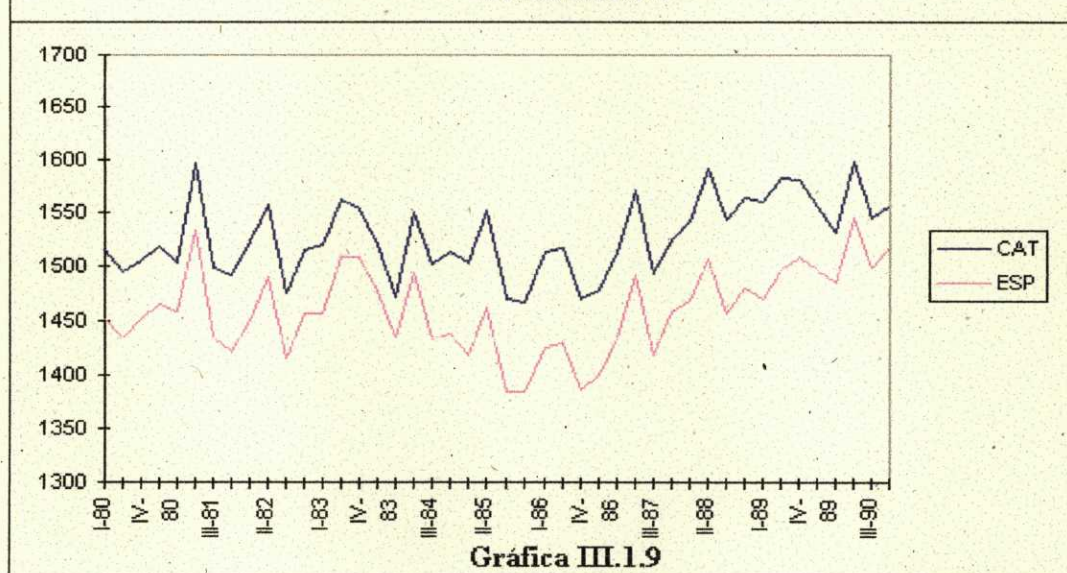
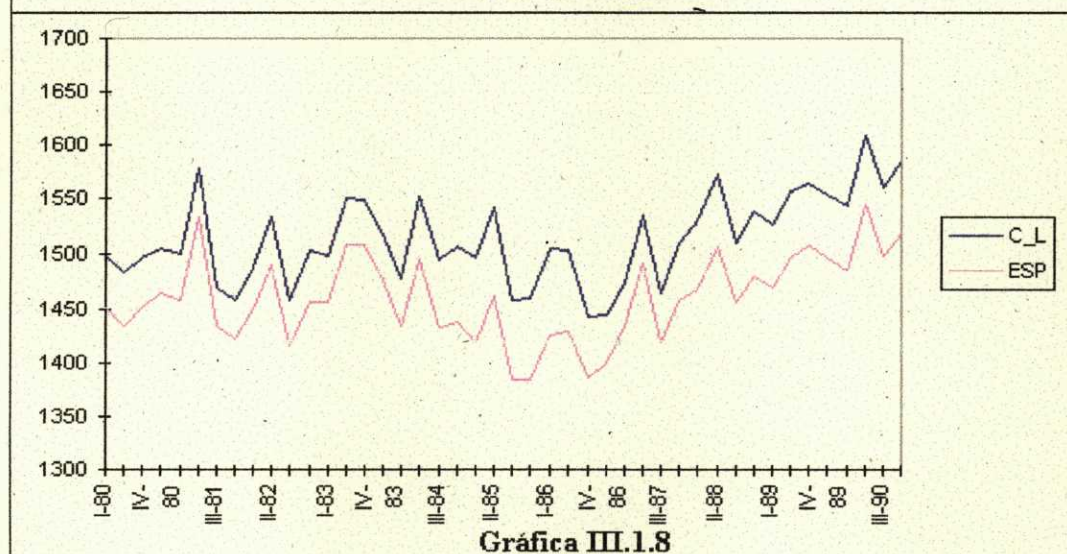
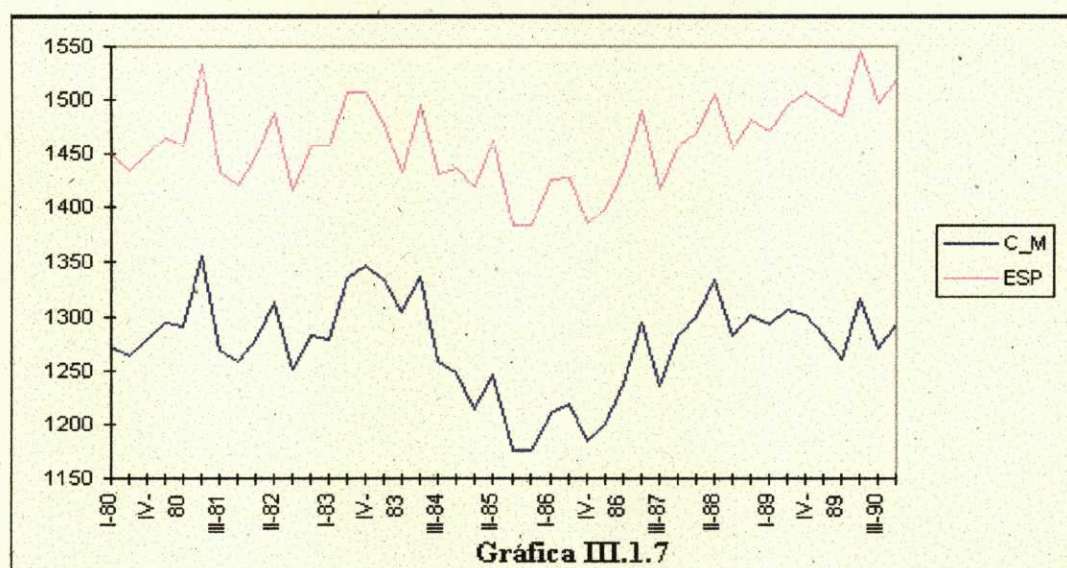


Gráfica III.1.3

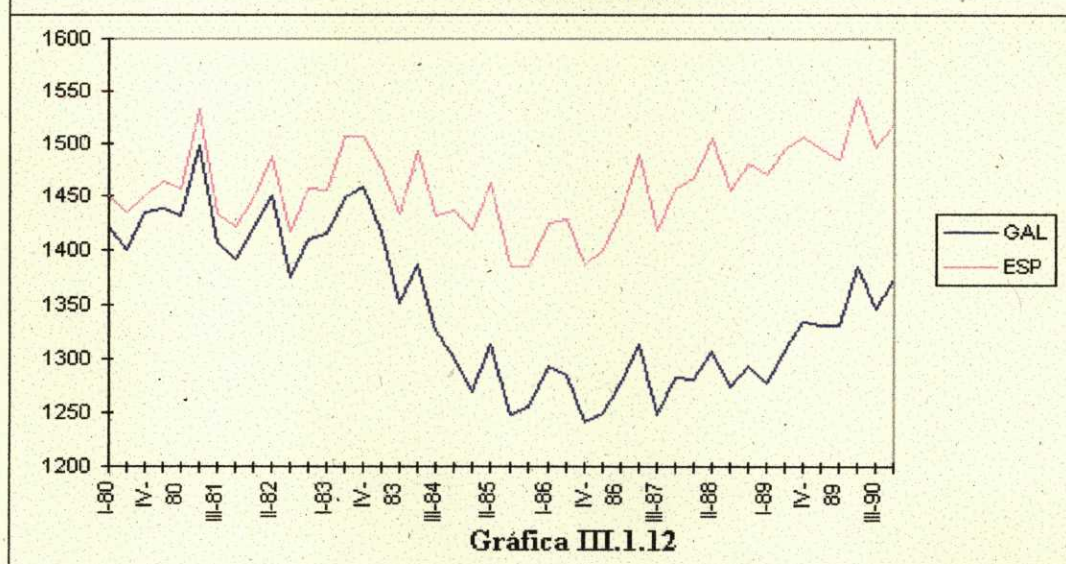
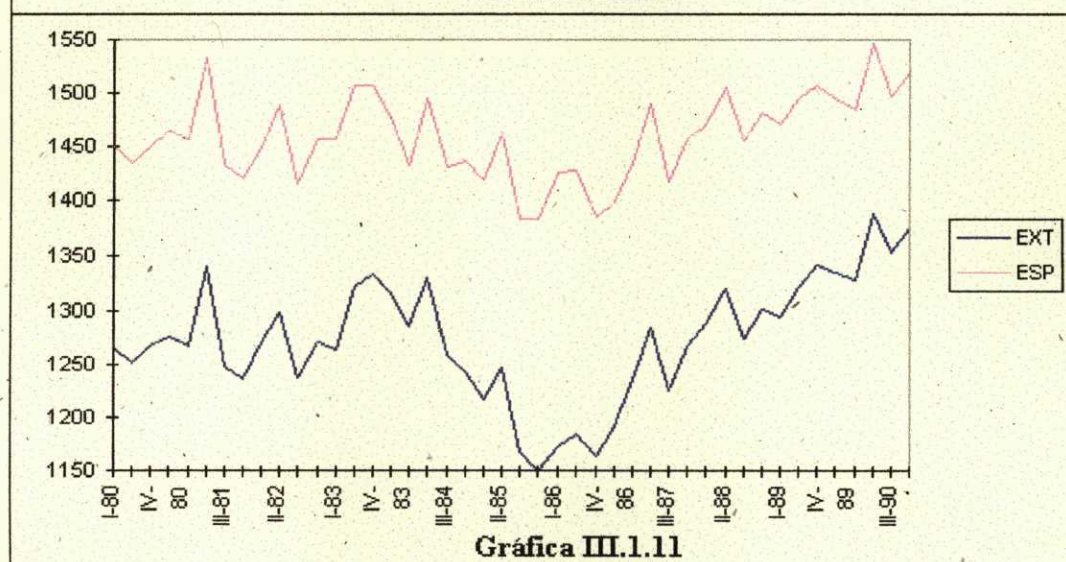
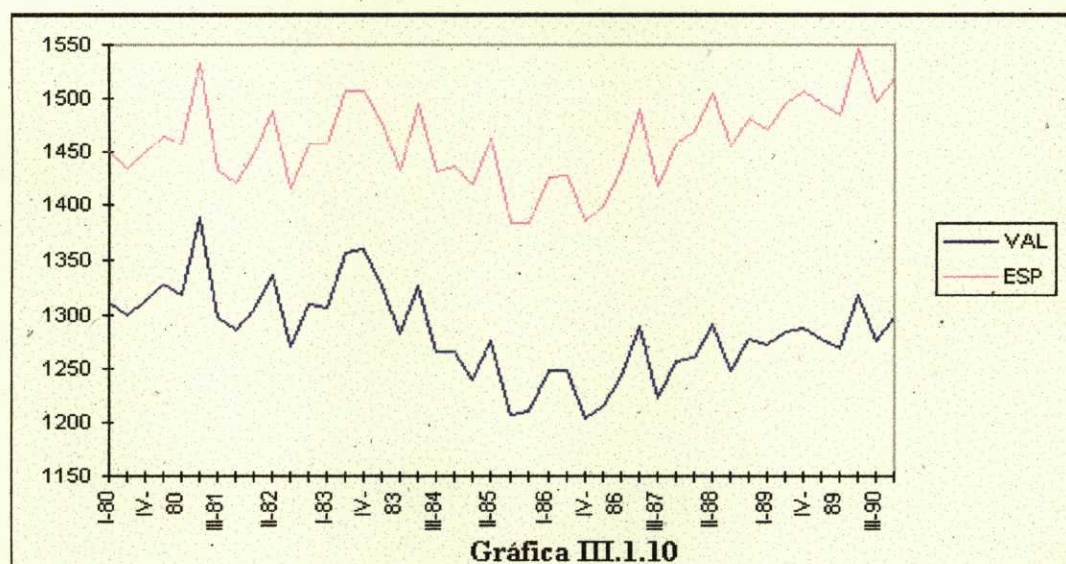
Nivel de Salarios. (Baleares, Canarias y Cantabria)



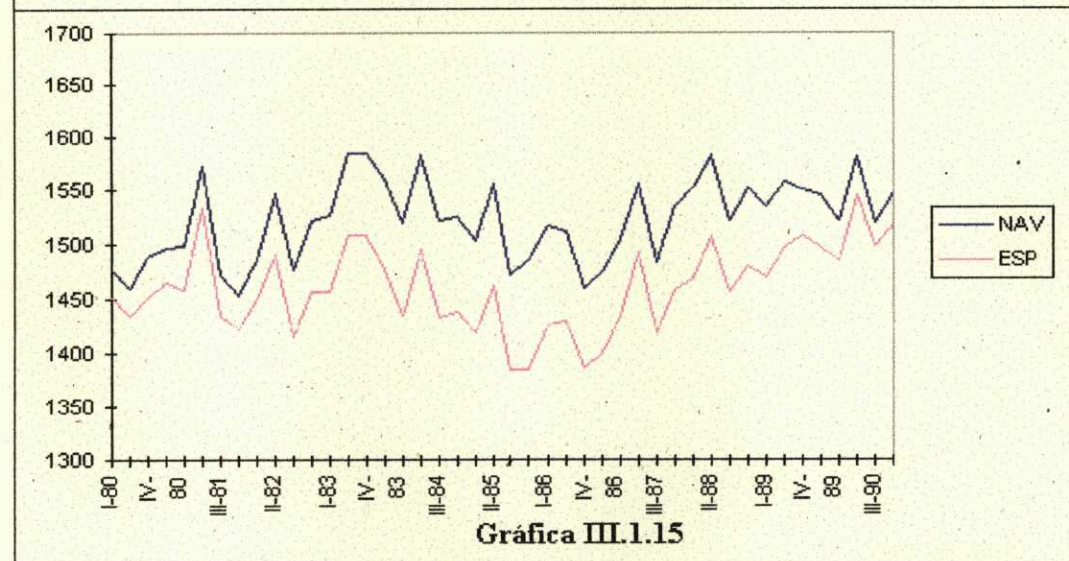
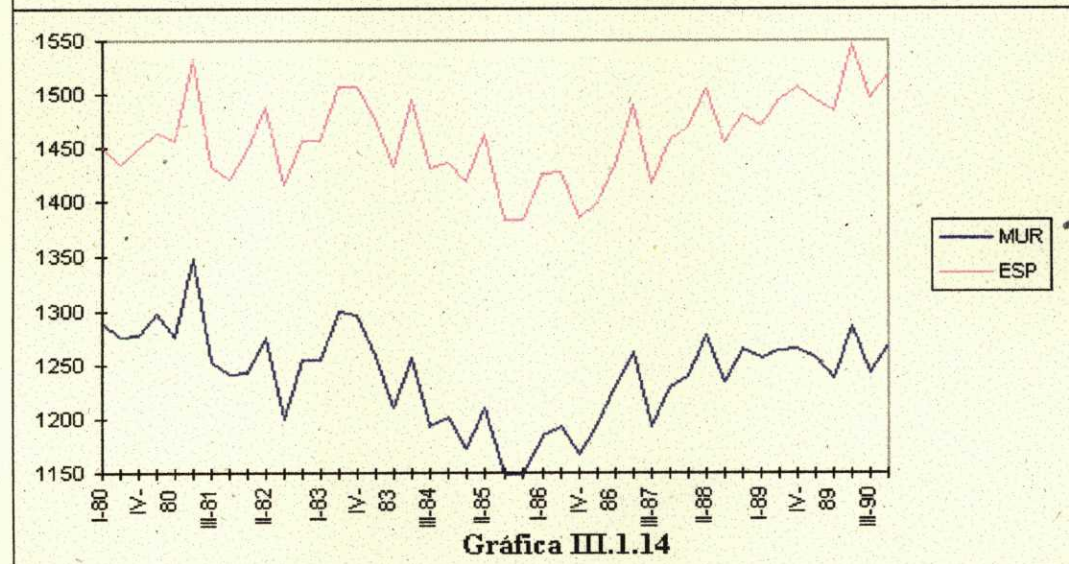
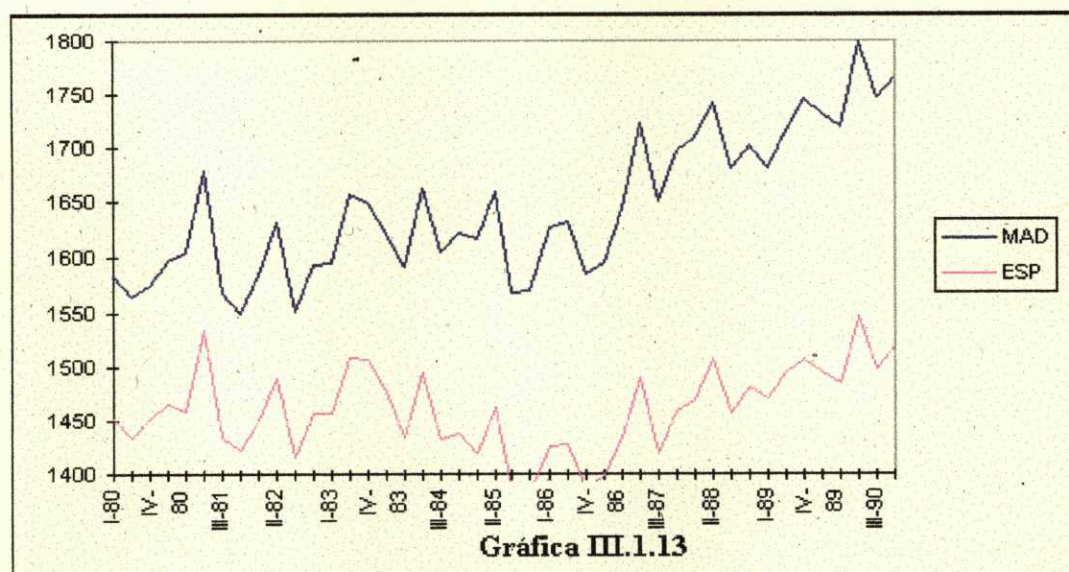
Nivel de Salarios. (Castilla - La Mancha, Castilla y León, y Cataluña)



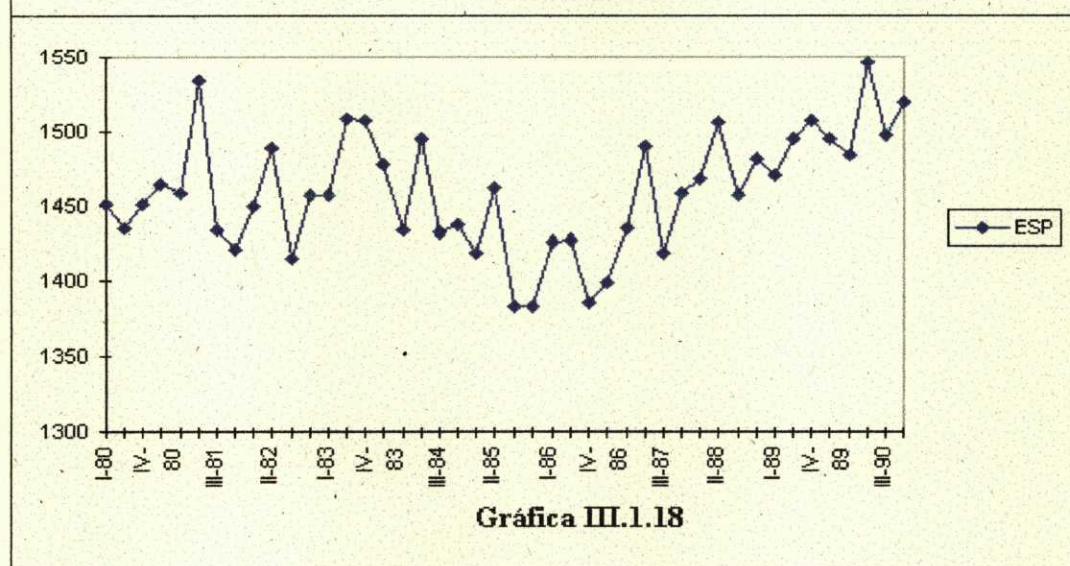
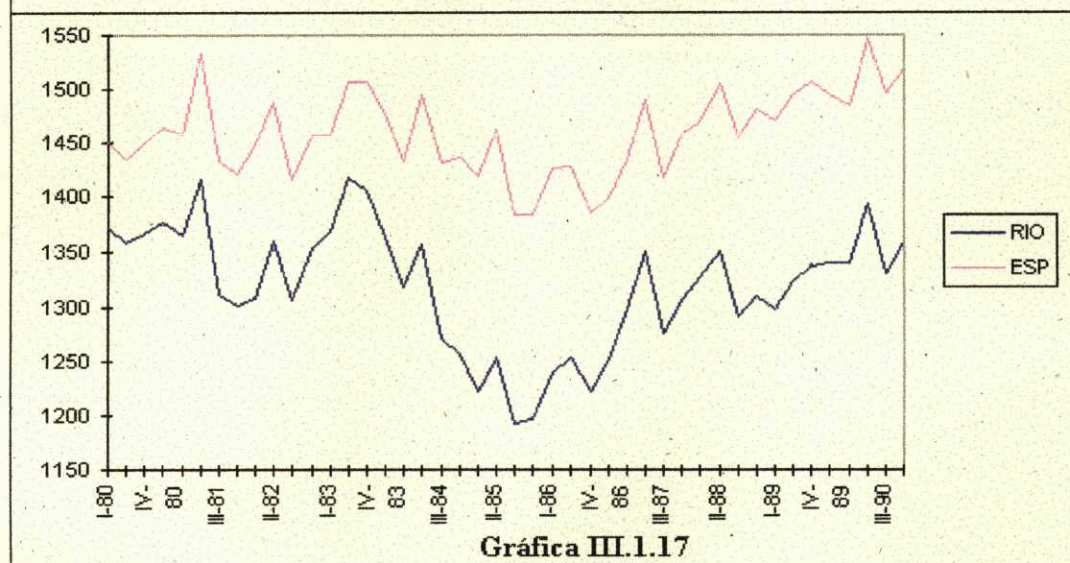
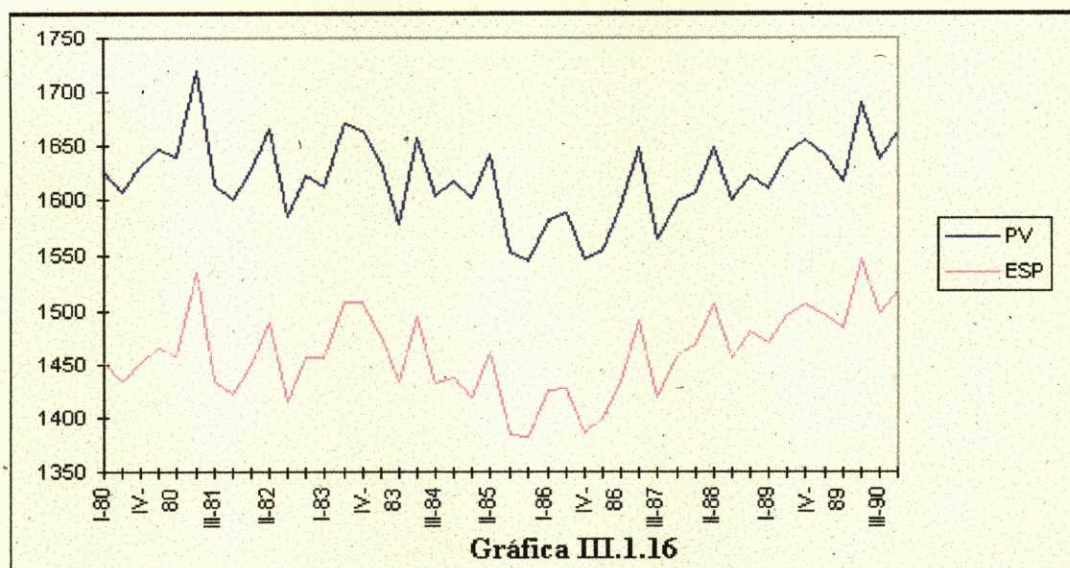
Nivel de Salarios. (Comunidad Valenciana, Extremadura y Galicia)



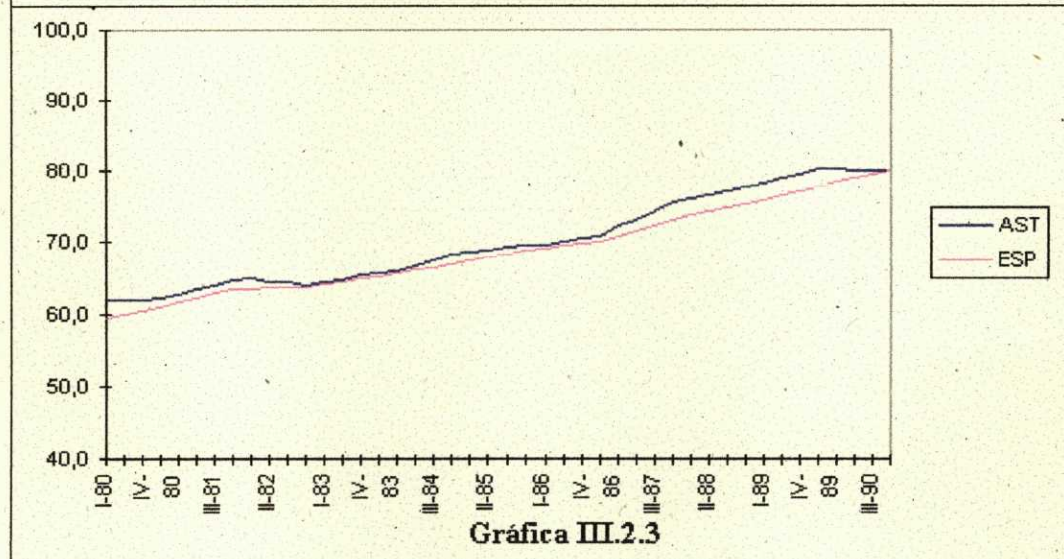
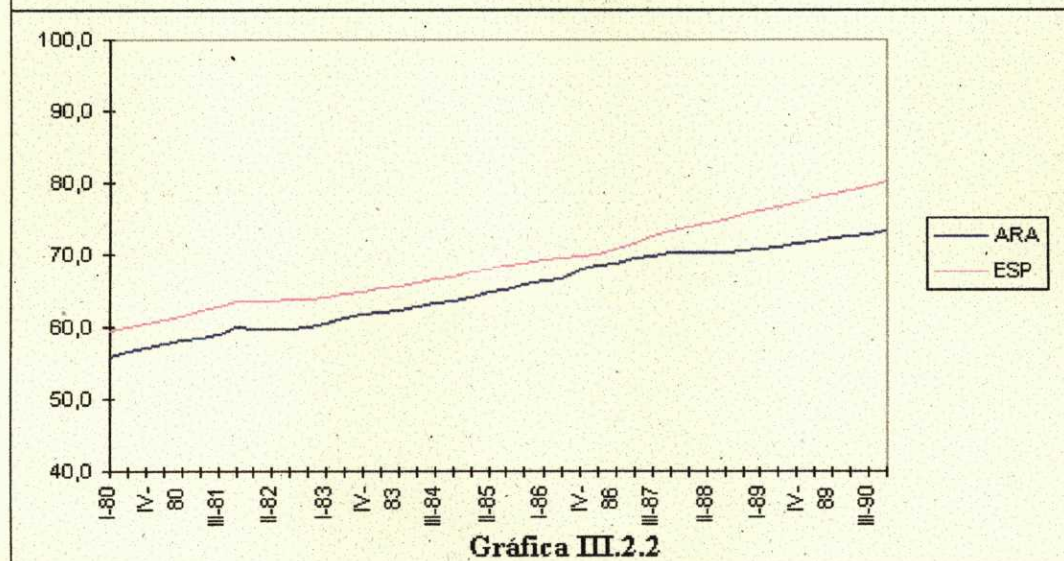
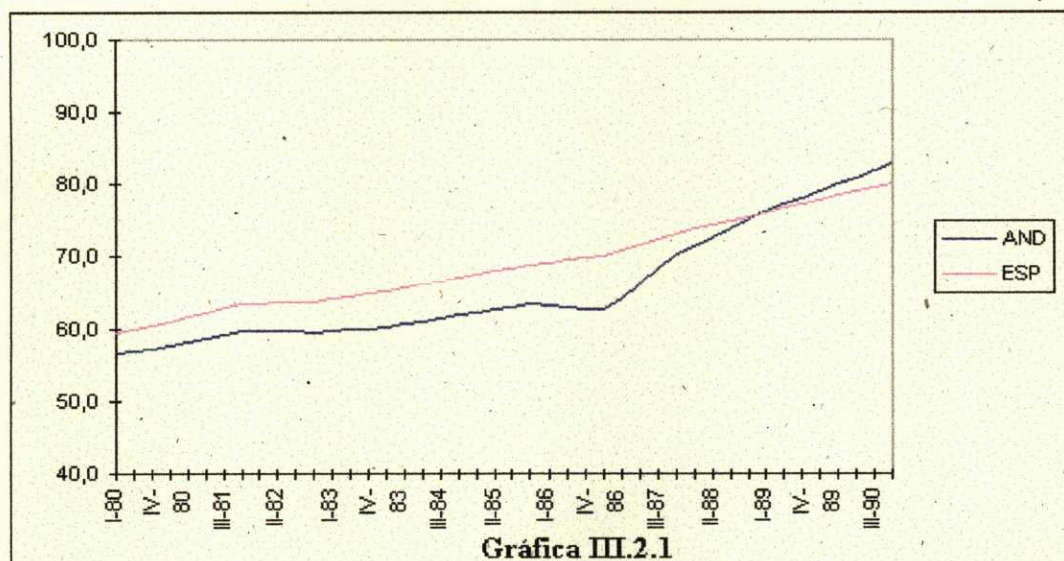
Nivel de Salarios. (Madrid, Murcia y Navarra)



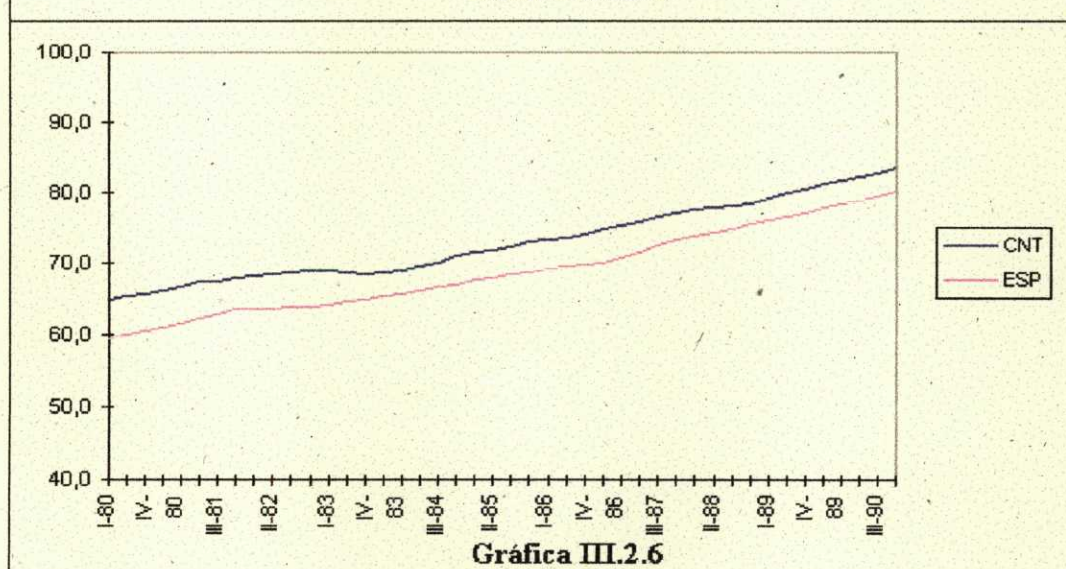
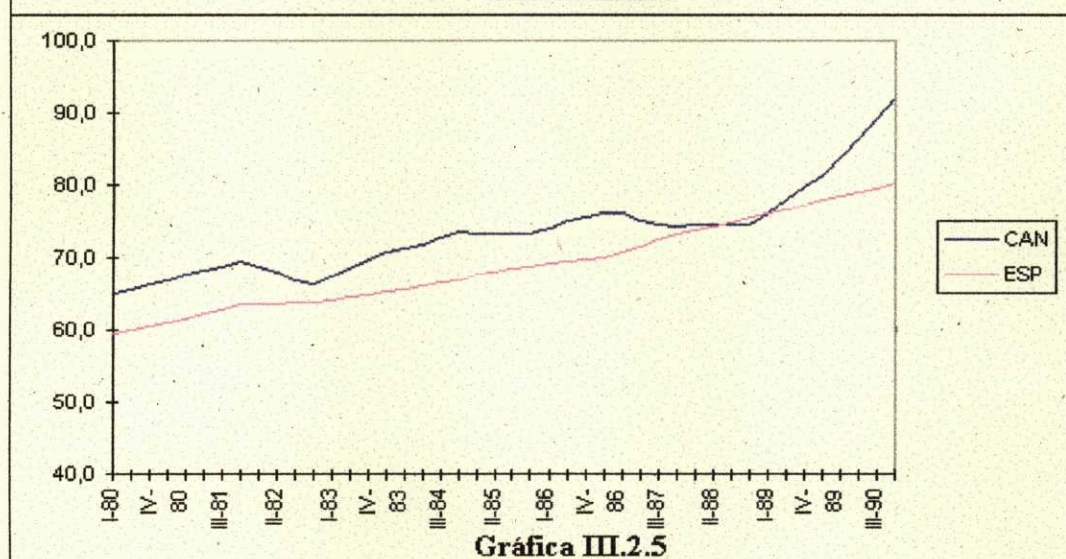
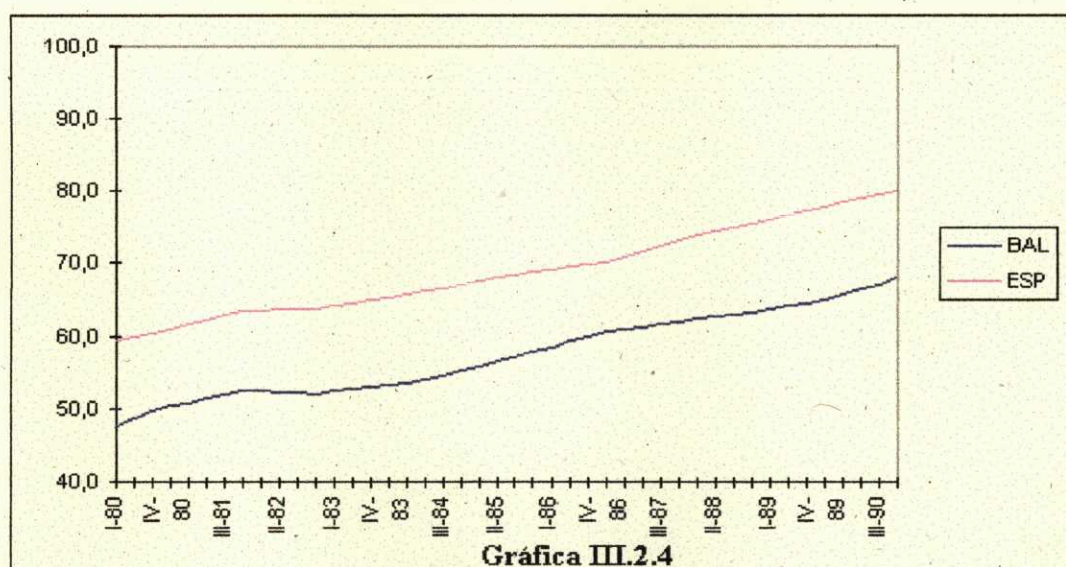
Nivel de Salarios. (País Vasco, La Rioja y datos globales)



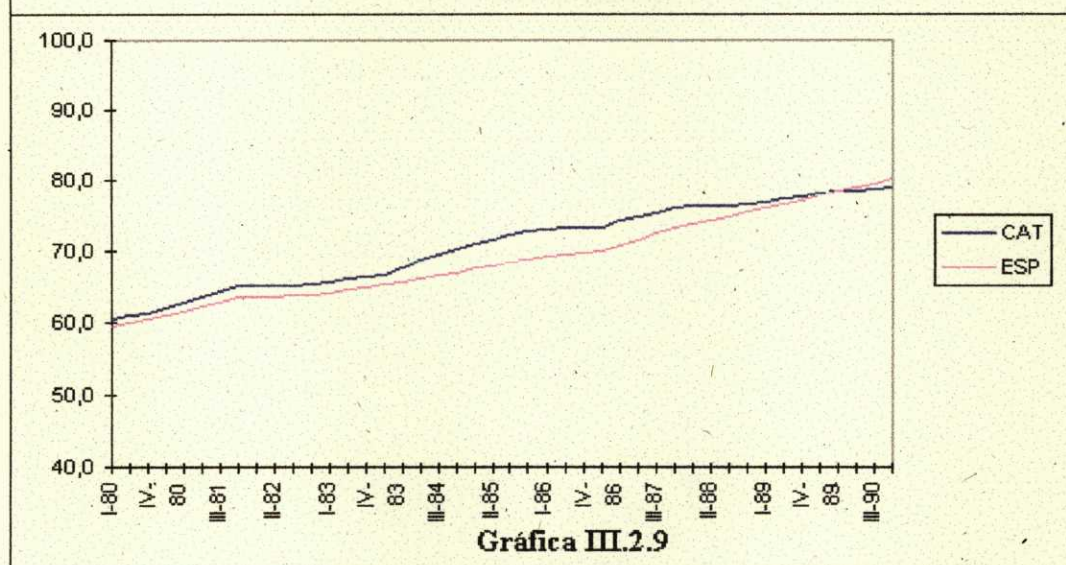
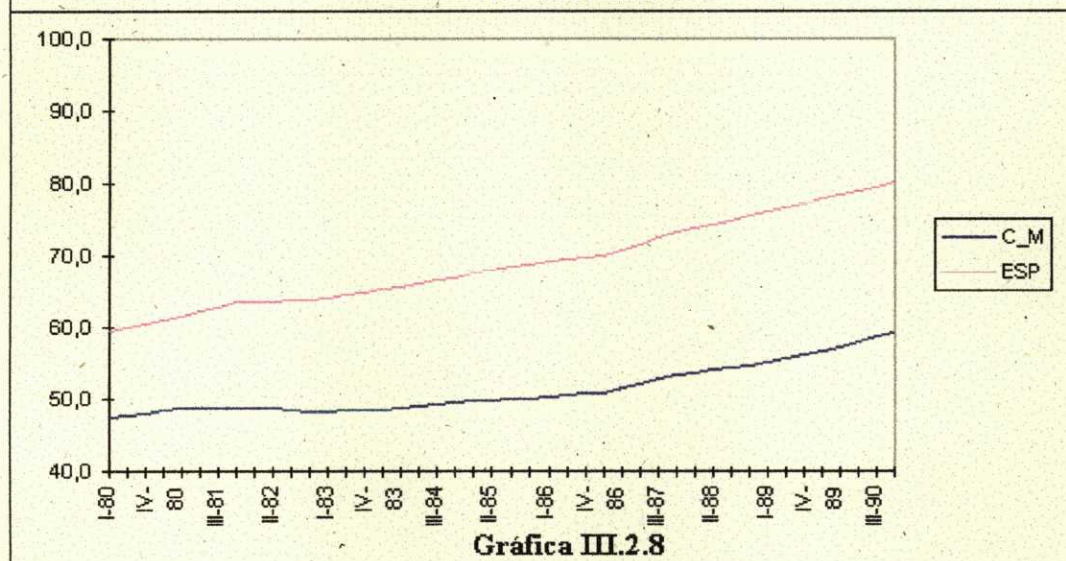
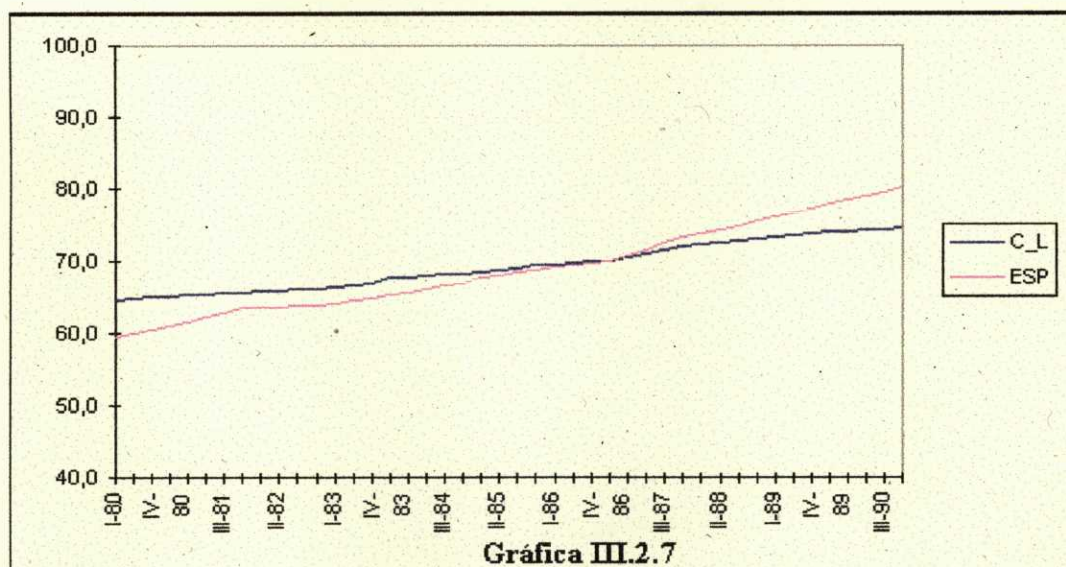
Nivel Educativo. (Andalucía, Aragón y Asturias)



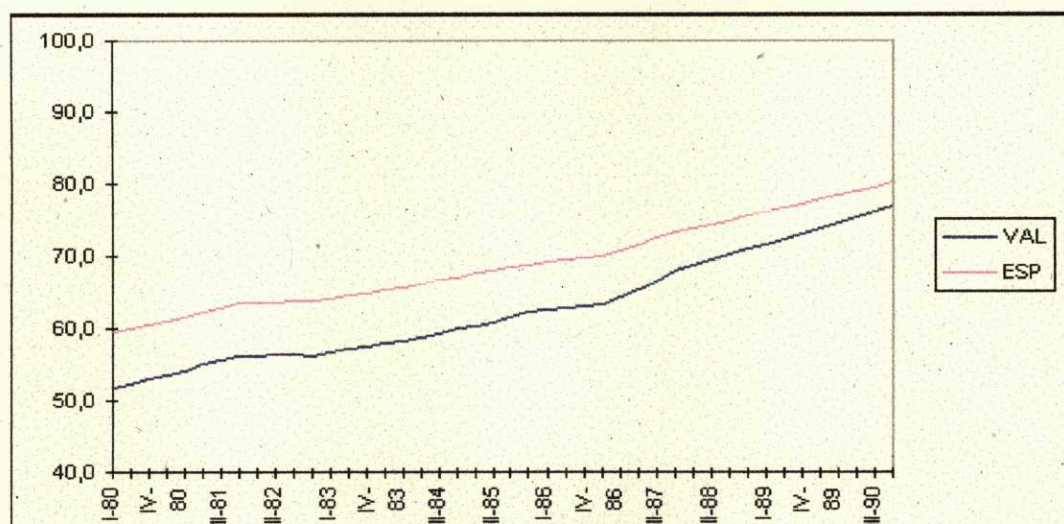
Nivel Educativo. (Baleares, Canarias y Cantabria)



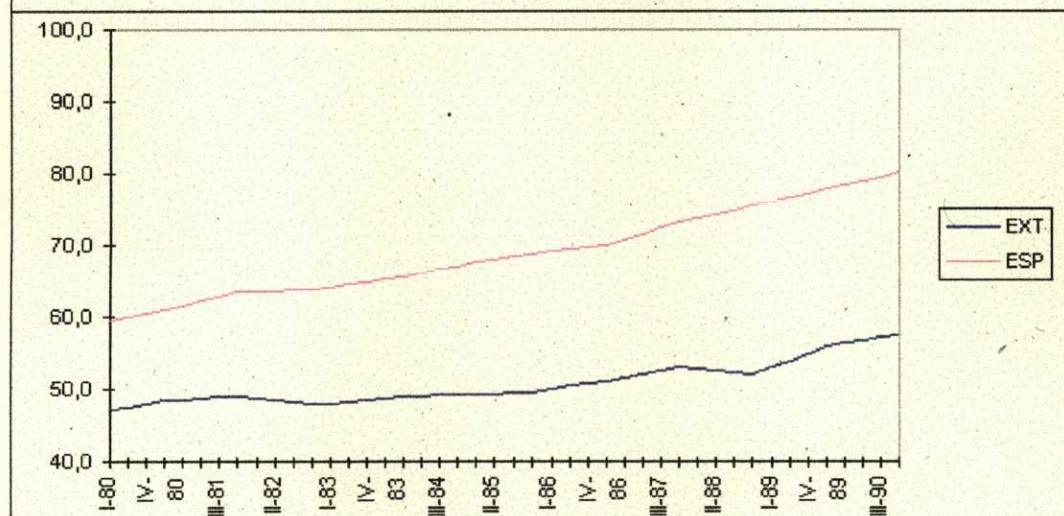
Nivel Educativo.
(Castilla - La Mancha, Castilla y León, y Cataluña)



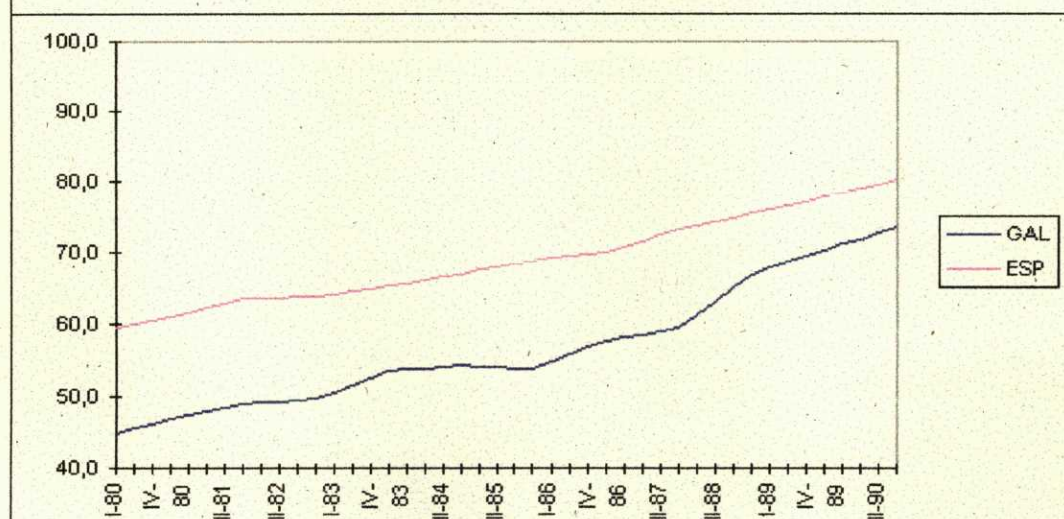
Nivel Educativo. (Comunidad Valenciana, Extremadura y Galicia)



Gráfica III.2.10

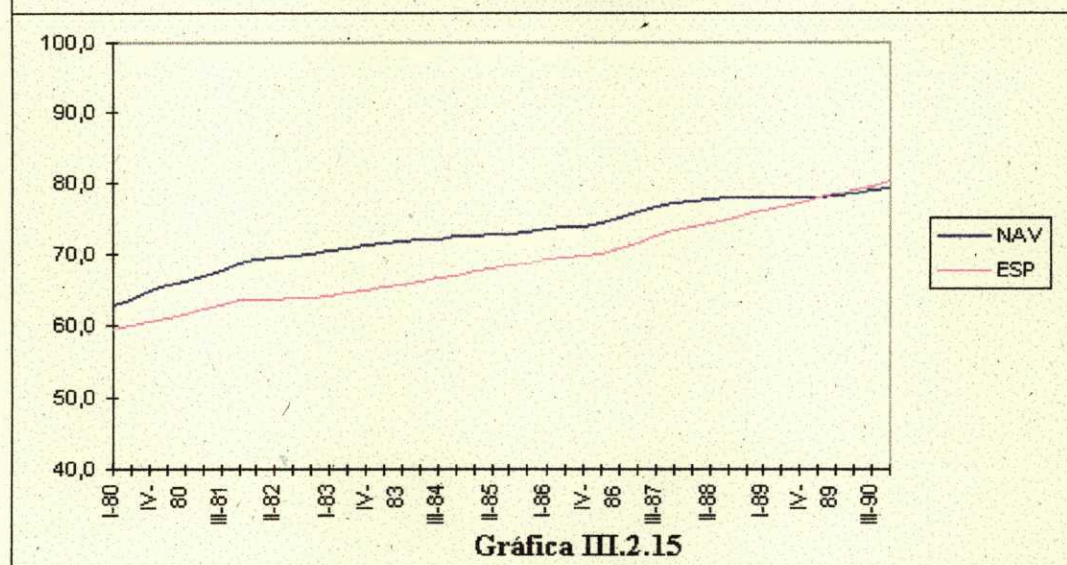
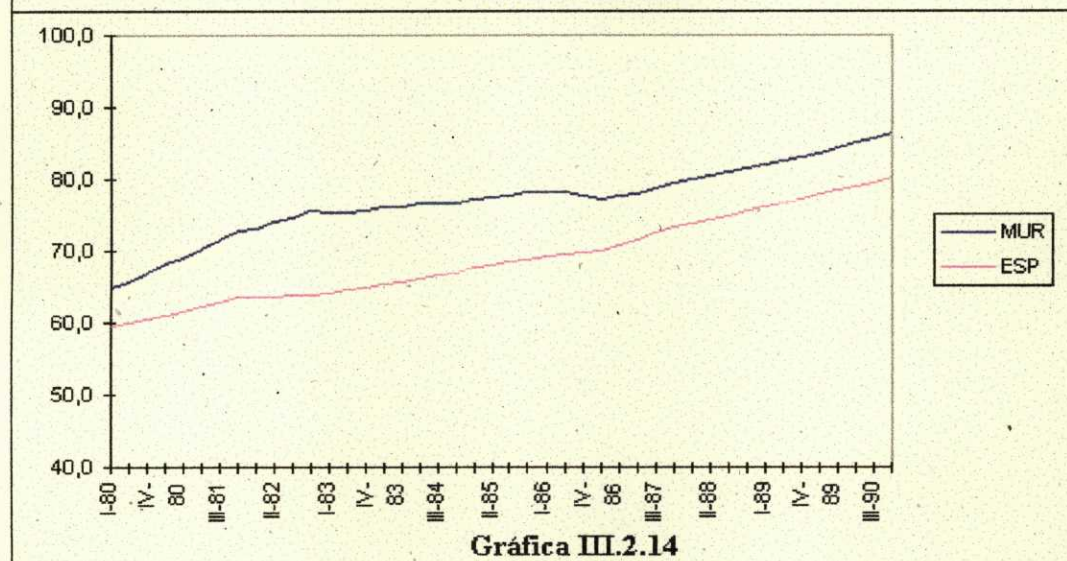
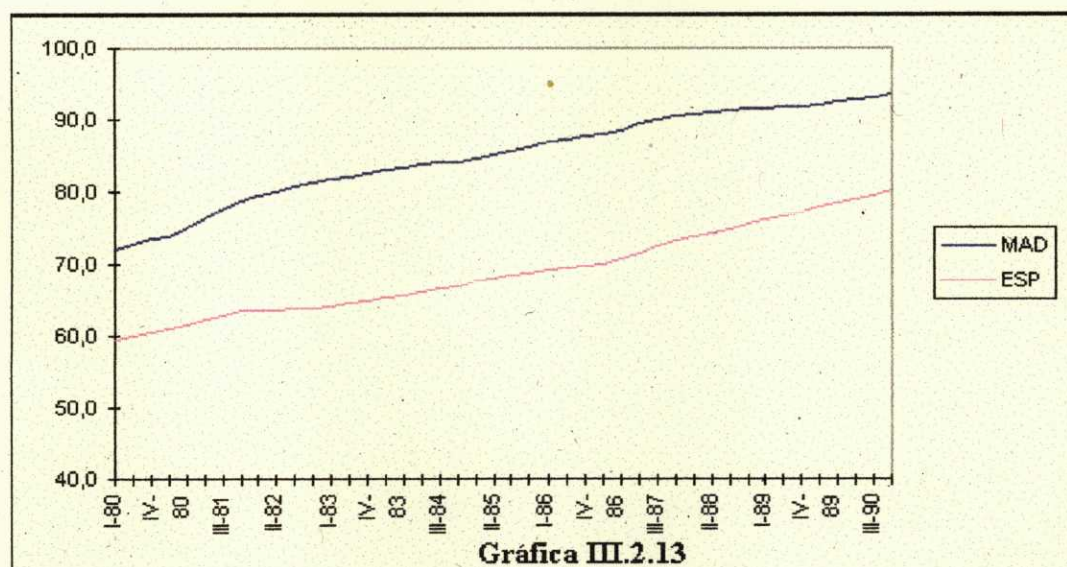


Gráfica III.2.11

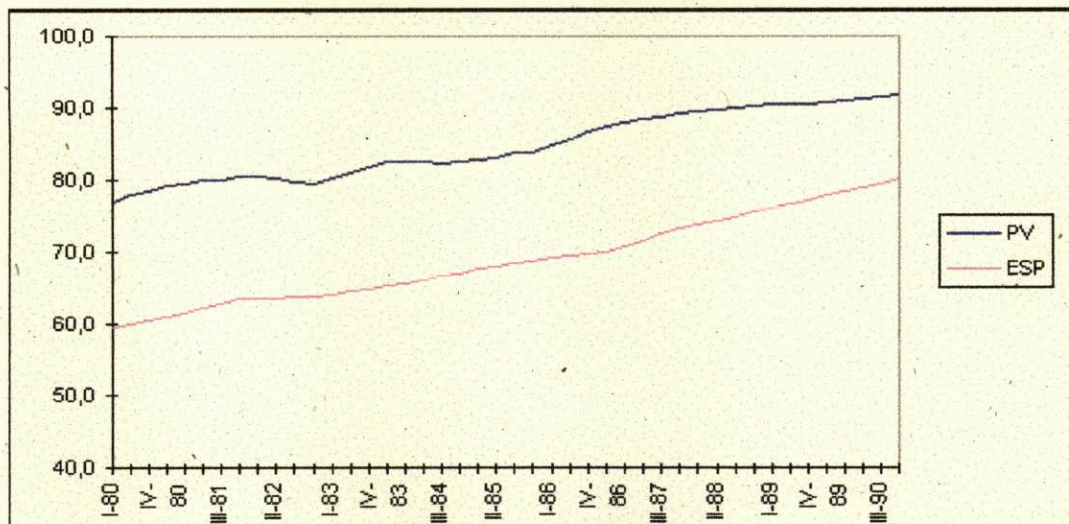


Gráfica III.2.12

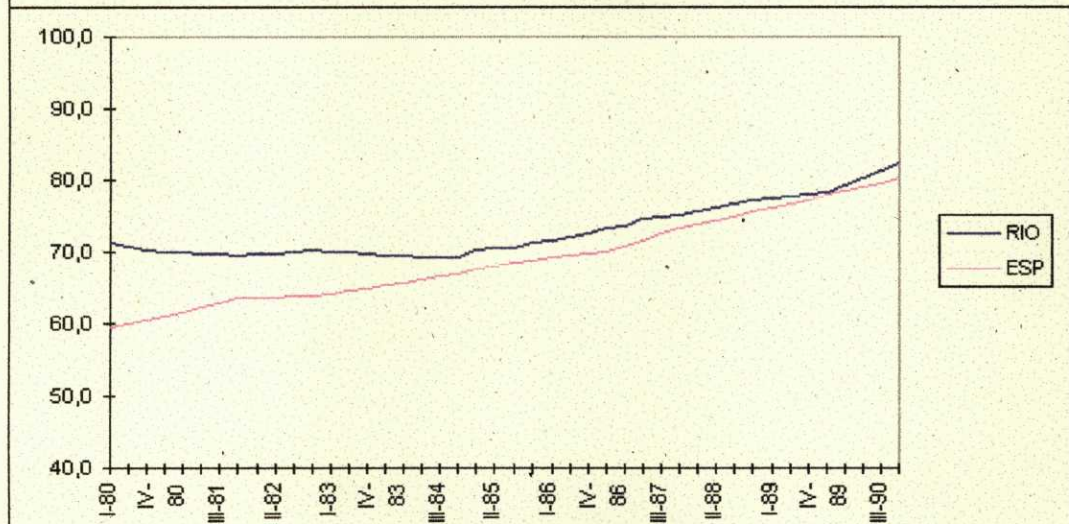
Nivel Educativo. (Madrid, Murcia y Navarra)



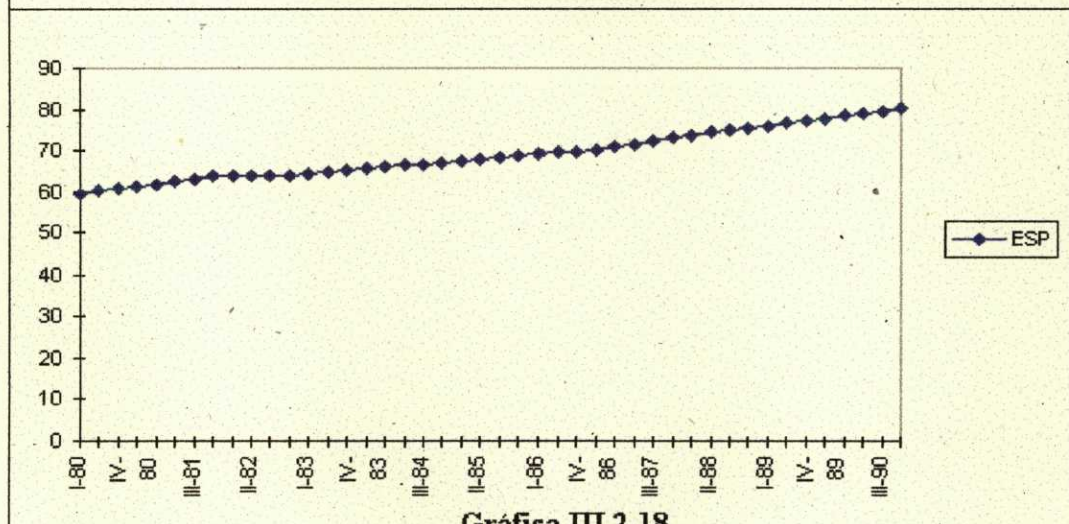
Nivel Educativo.
(País Vasco, La Rioja y datos globales)



Gráfica III.2.16

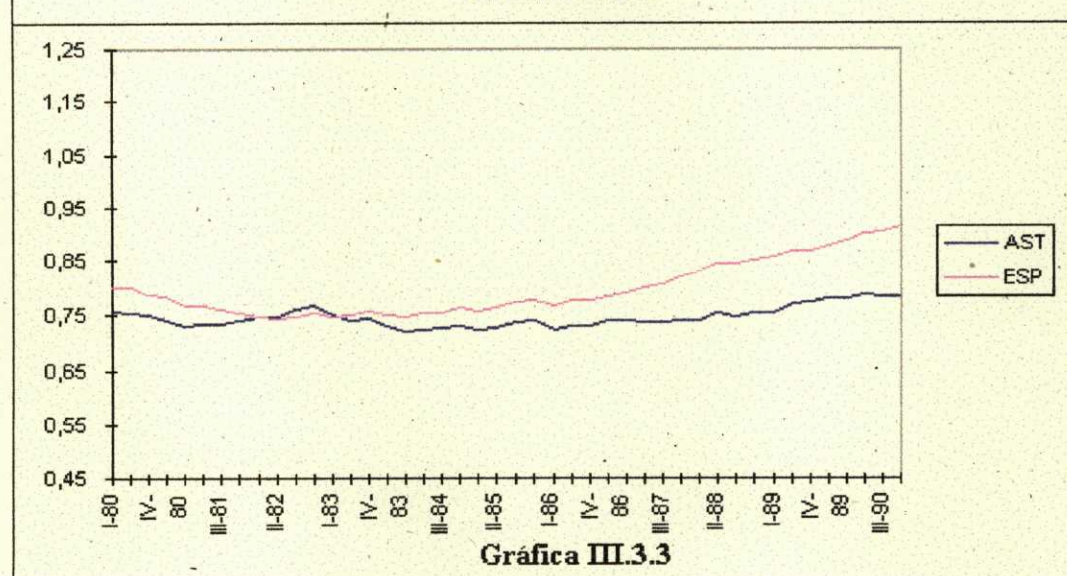
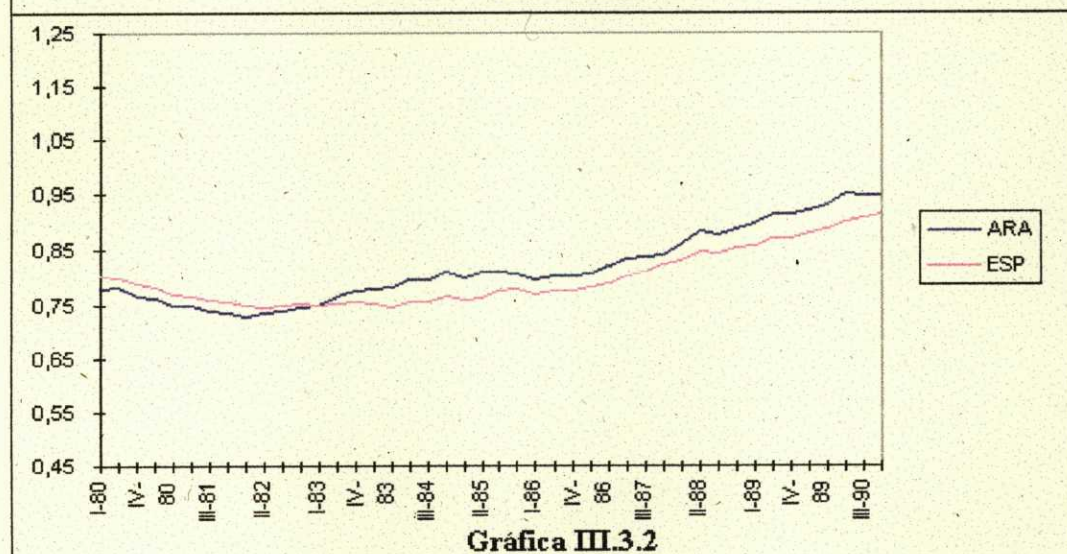
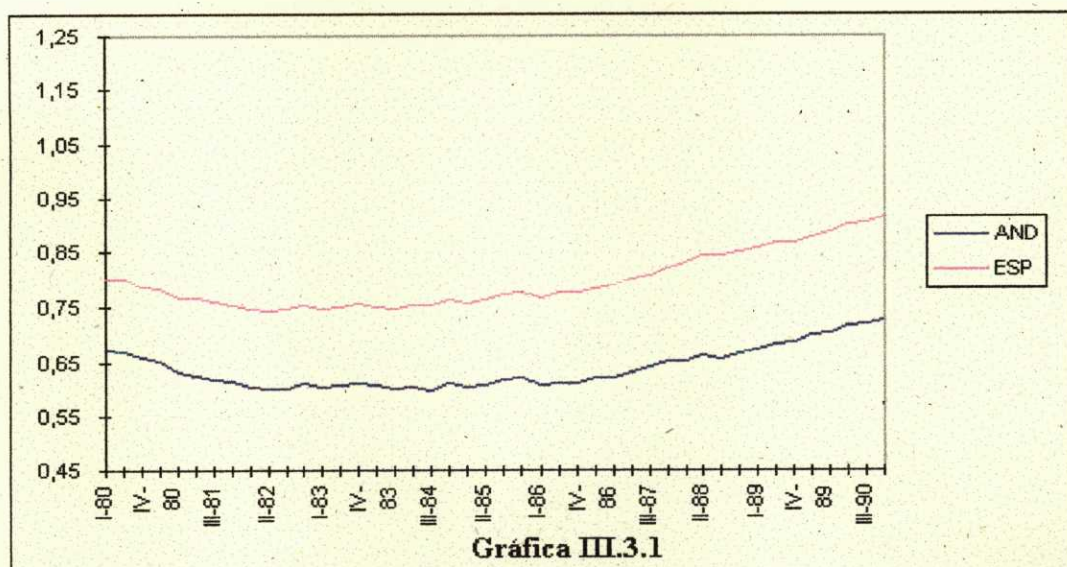


Gráfica III.2.17

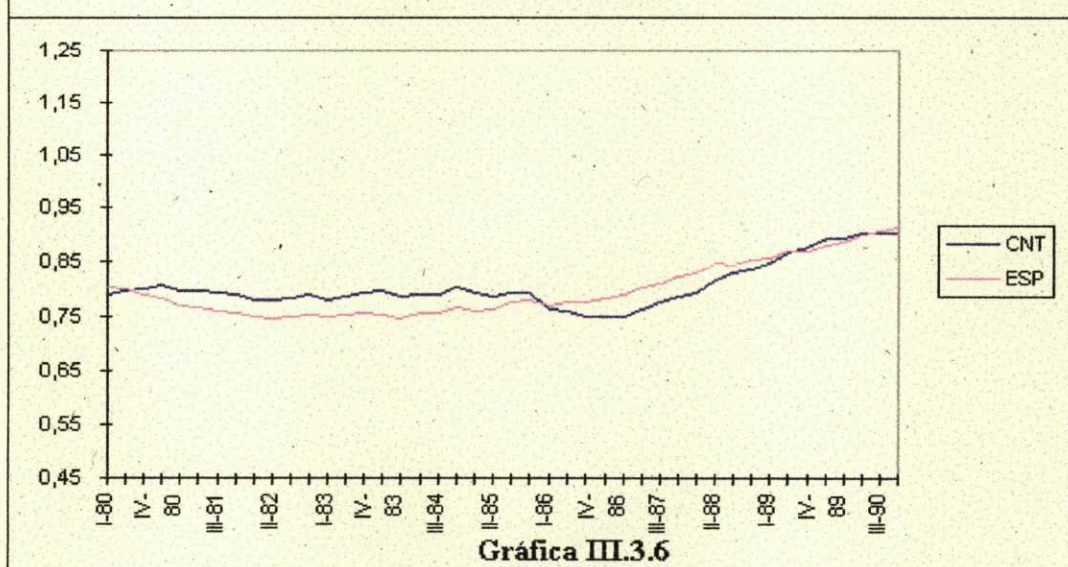
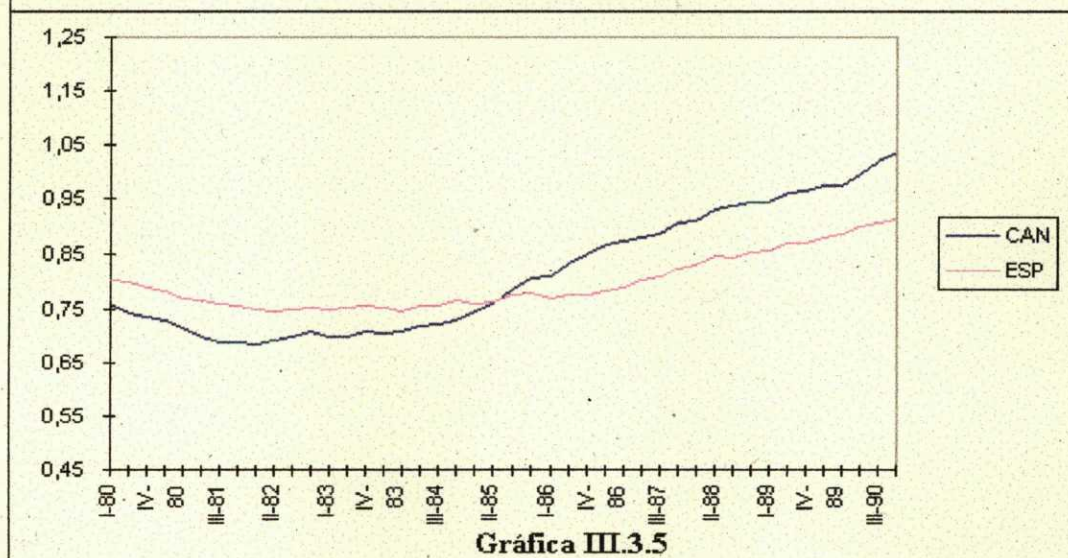
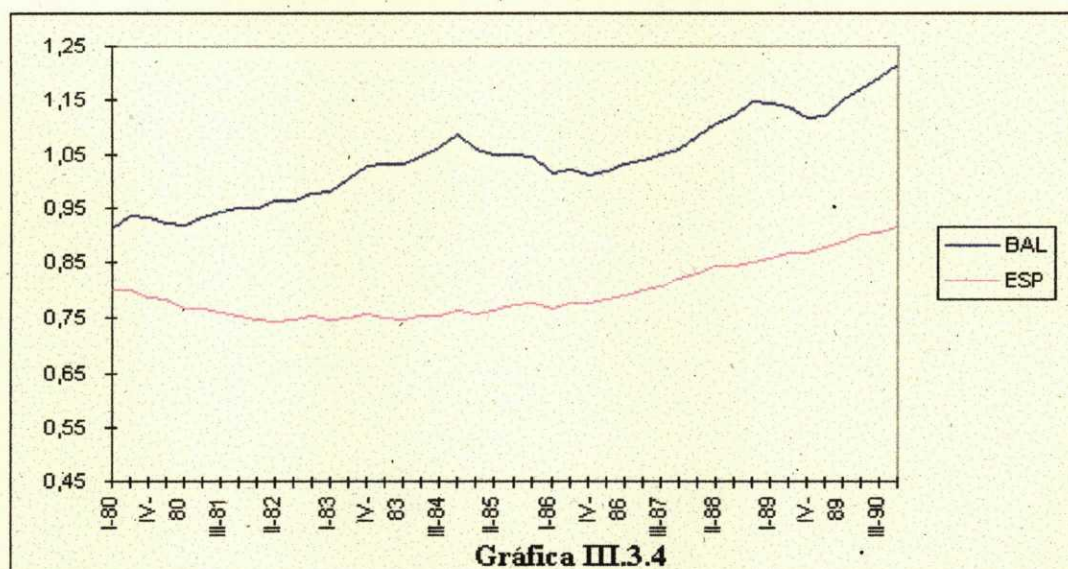


Gráfica III.2.18

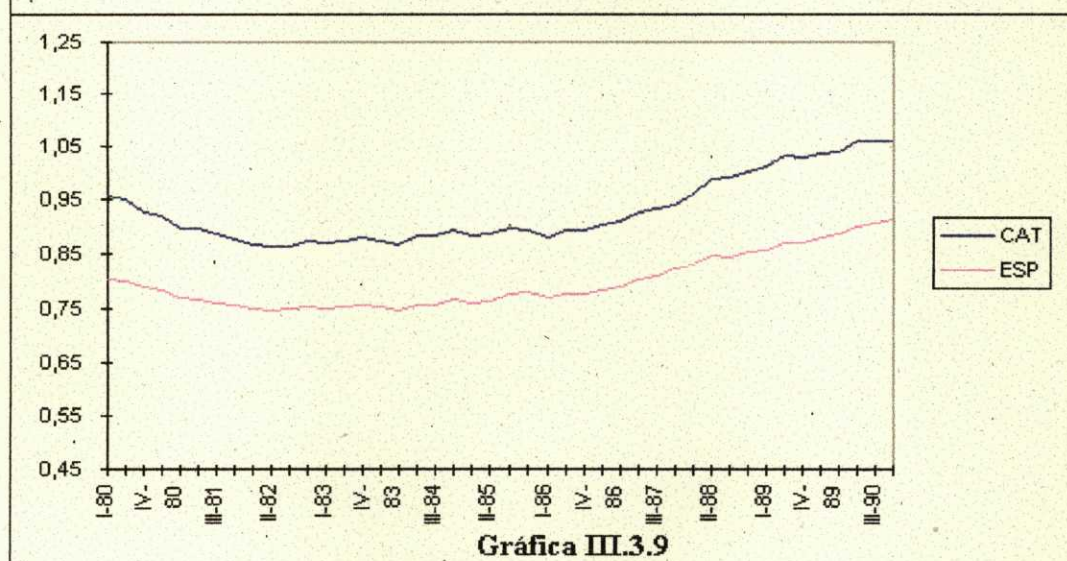
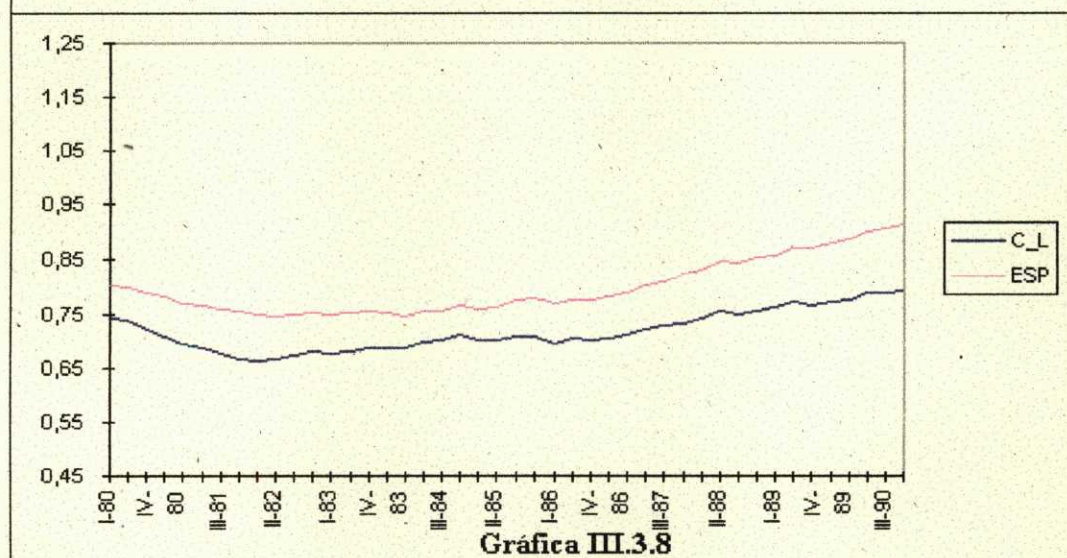
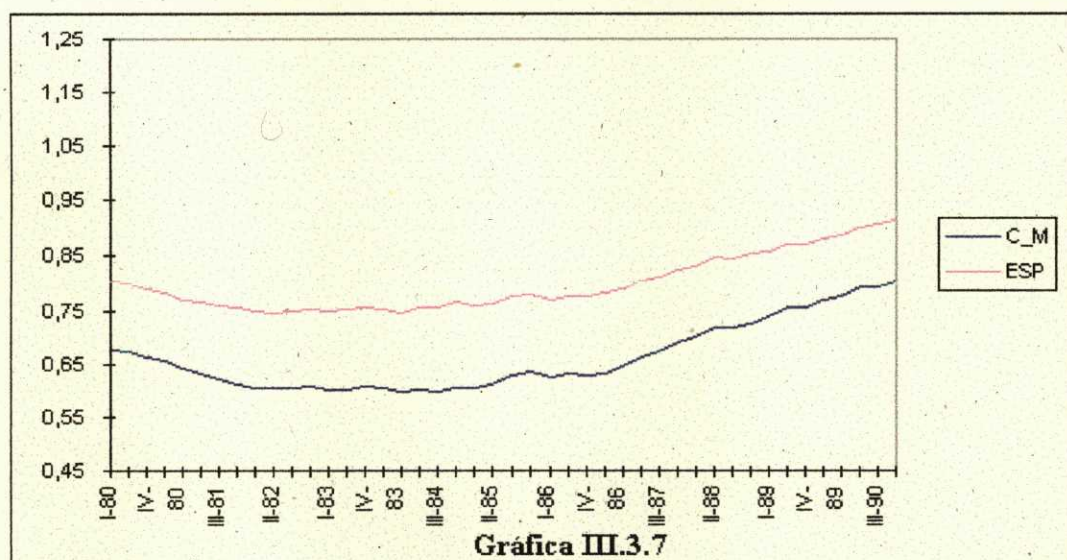
P.I.B. por Habitante. (Andalucía, Aragón y Asturias)



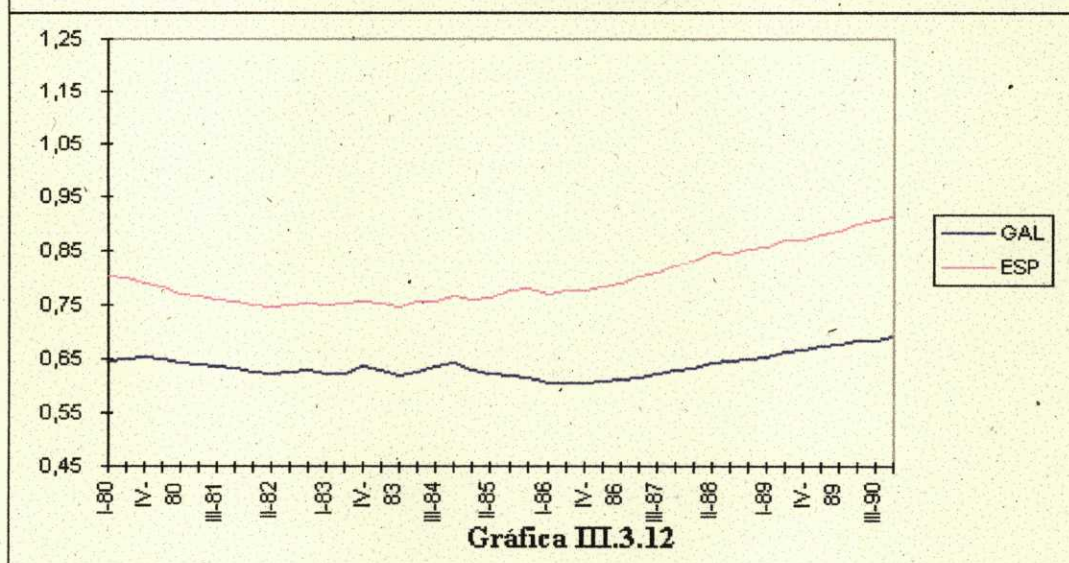
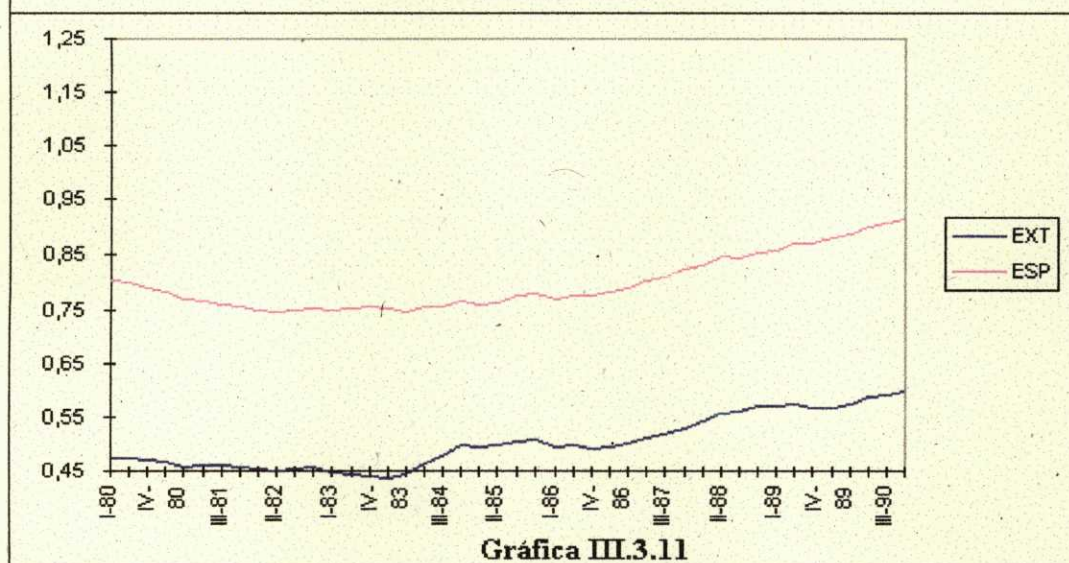
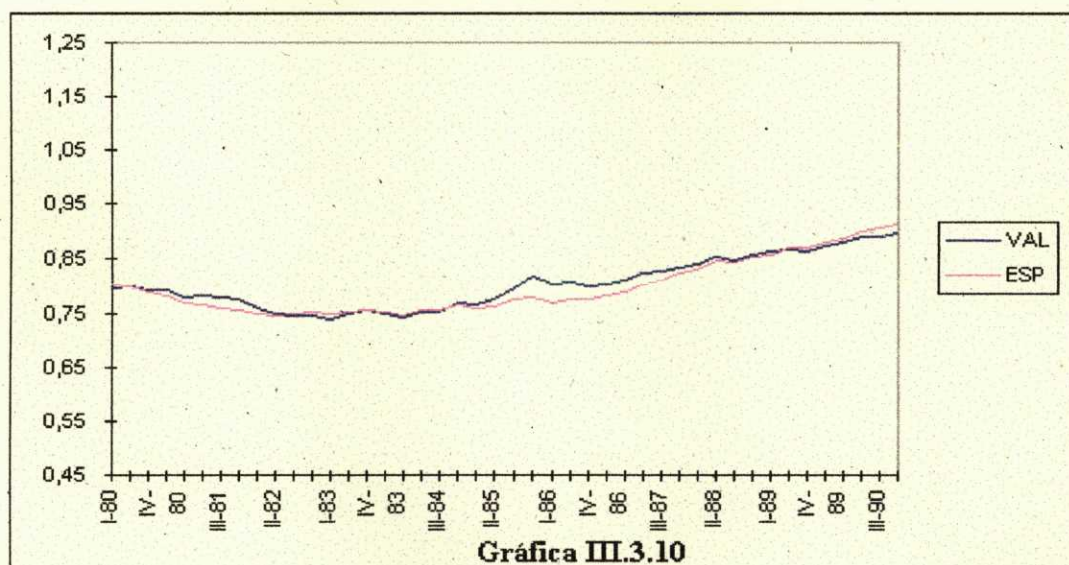
P.I.B. por Habitante.
(Baleares, Canarias y Cantabria)



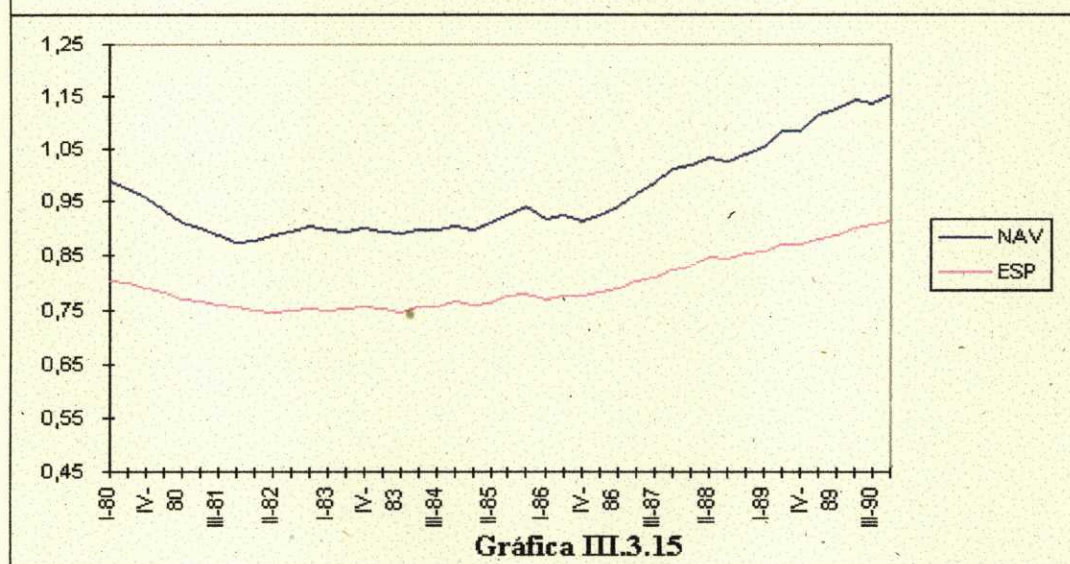
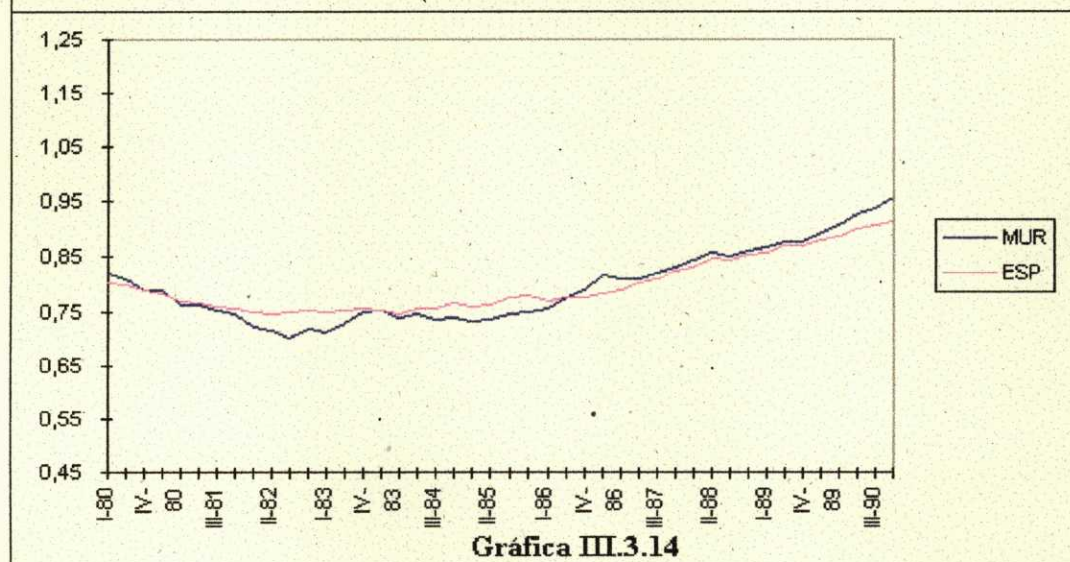
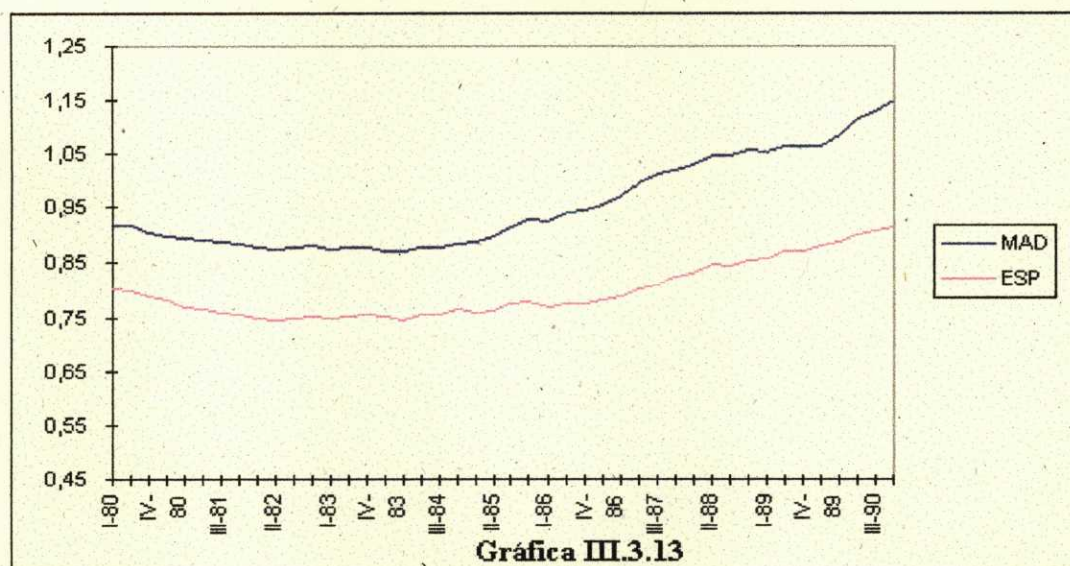
P.I.B. por Habitante.
(Castilla - La Mancha, Castilla y León, y Cataluña)



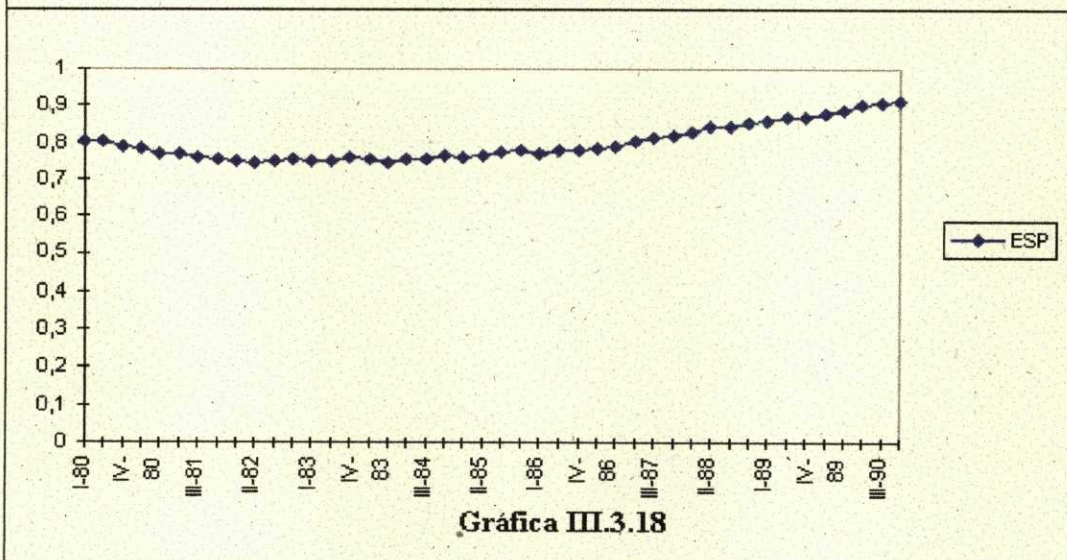
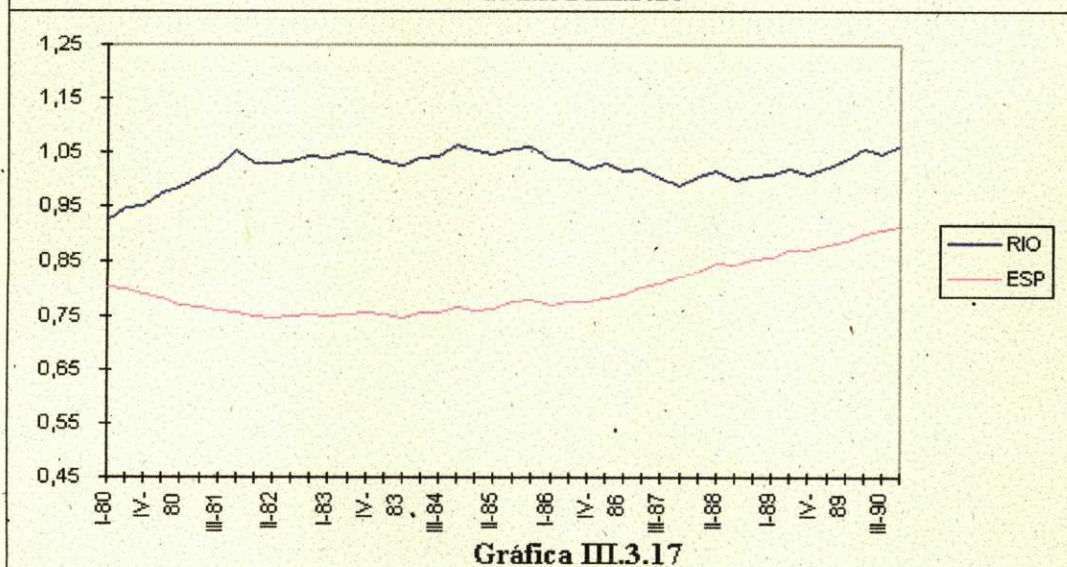
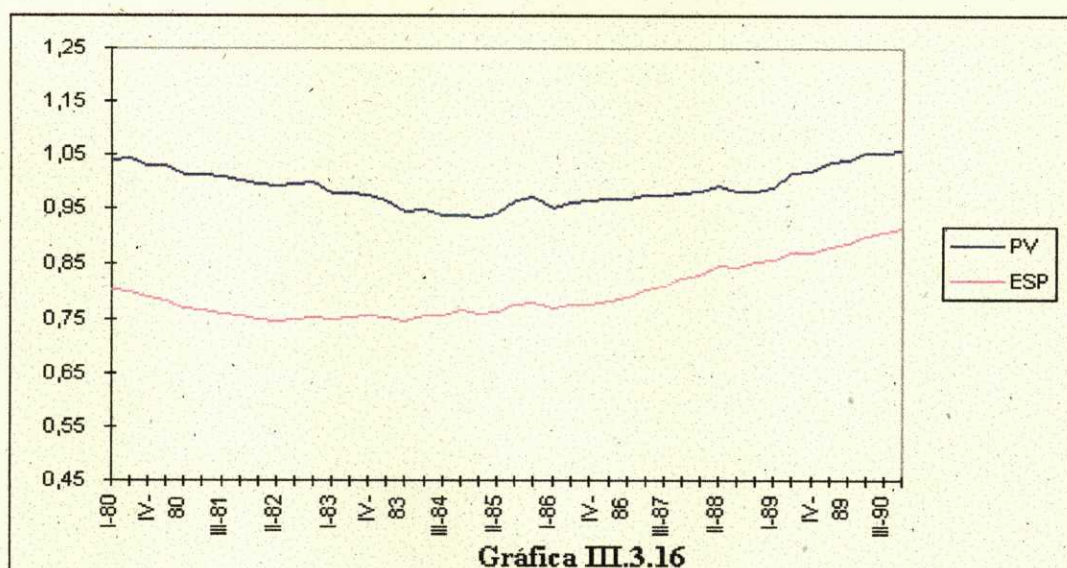
P.I.B. por Habitante.
(Comunidad Valenciana, Extremadura y Galicia)



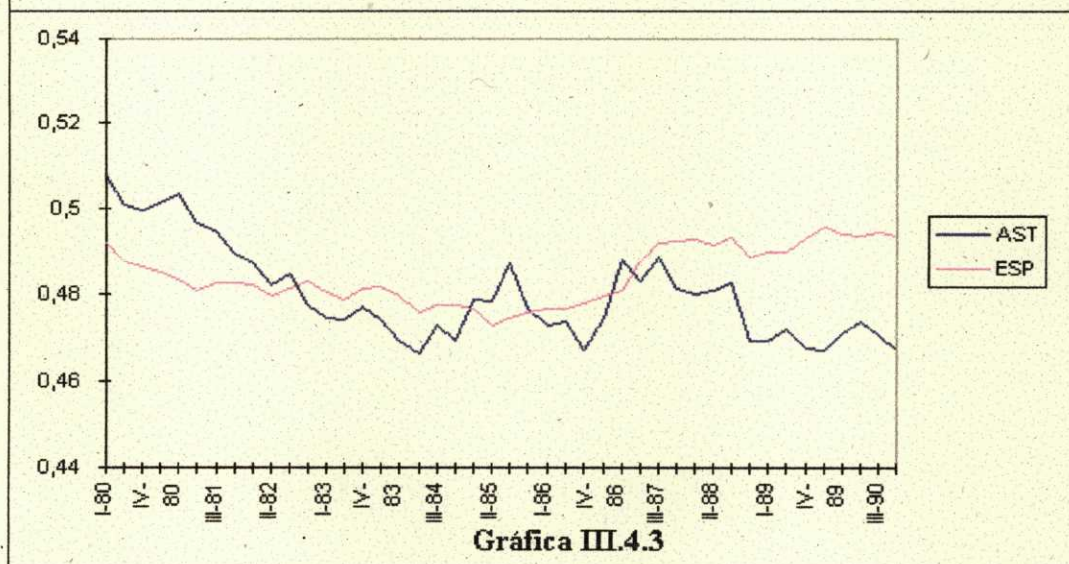
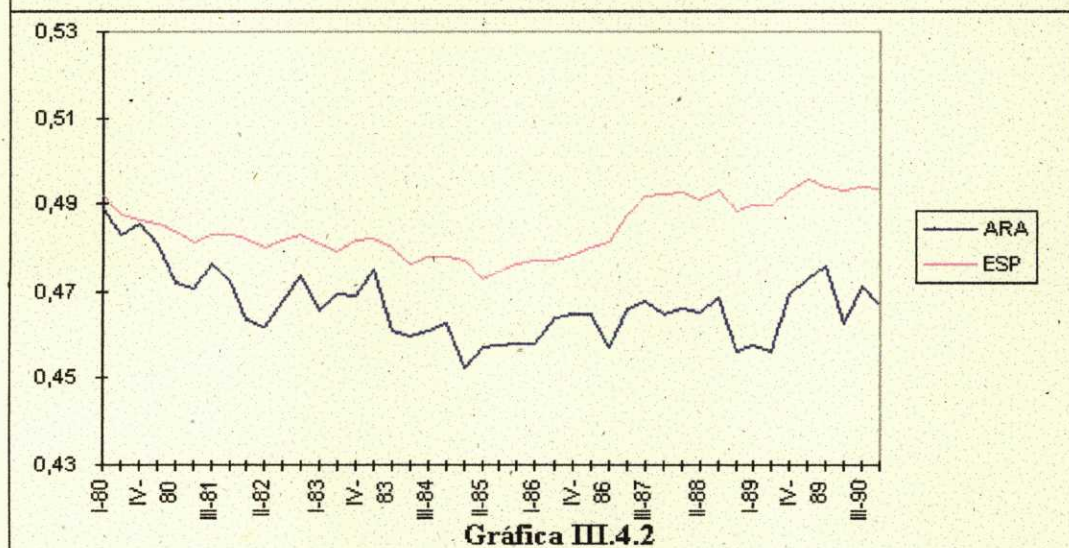
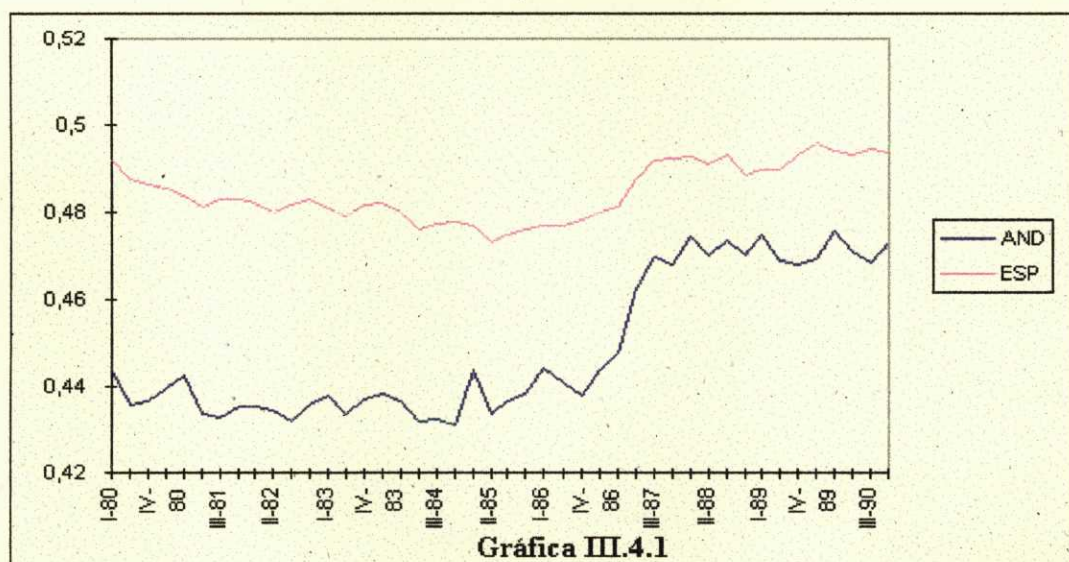
P.I.B. por Habitante. (Madrid, Murcia y Navarra)



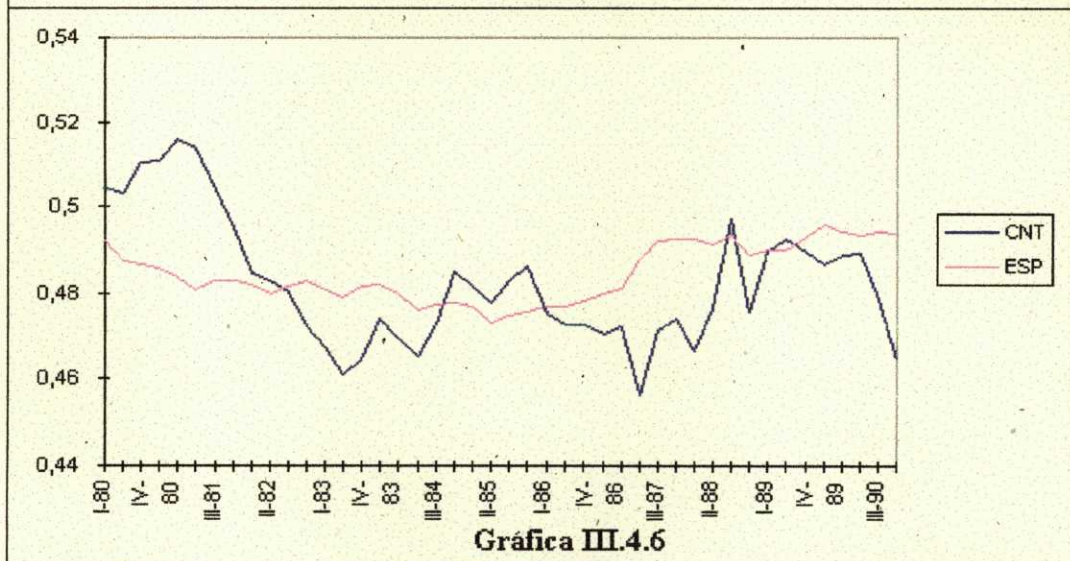
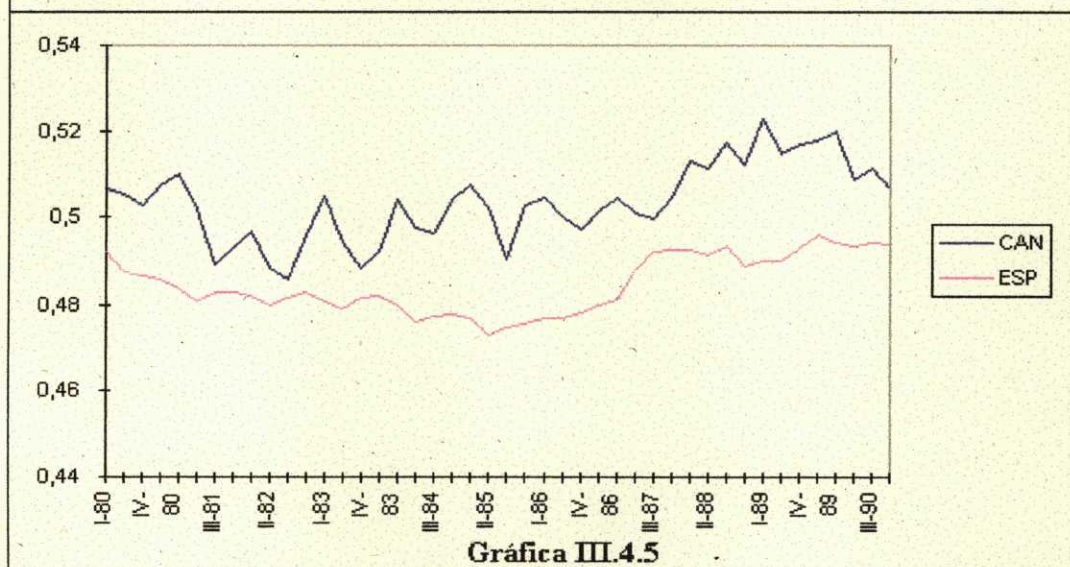
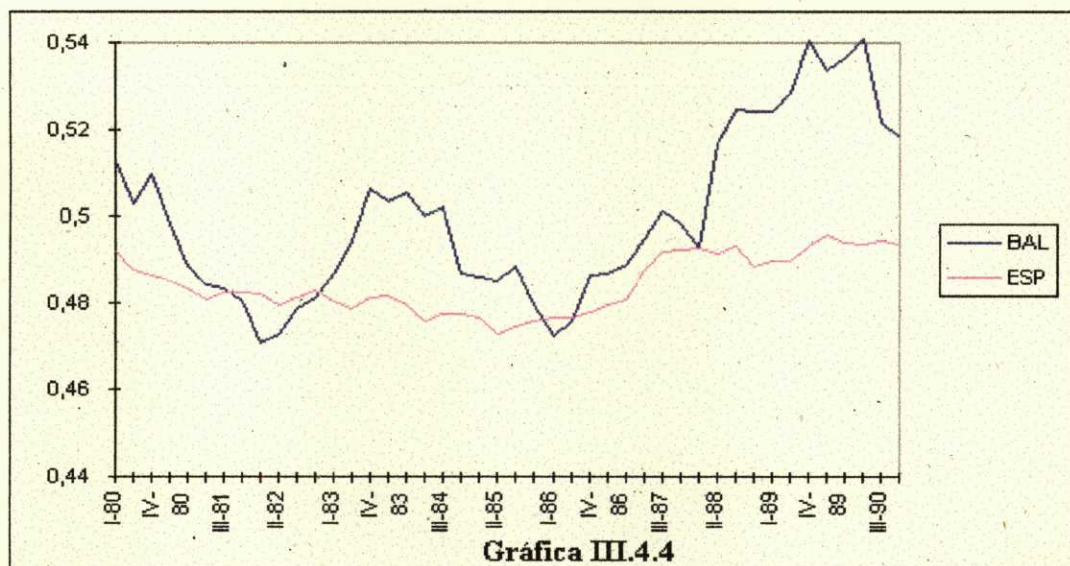
P.I.B. por Habitante.
(País Vasco, La Rioja y datos globales)



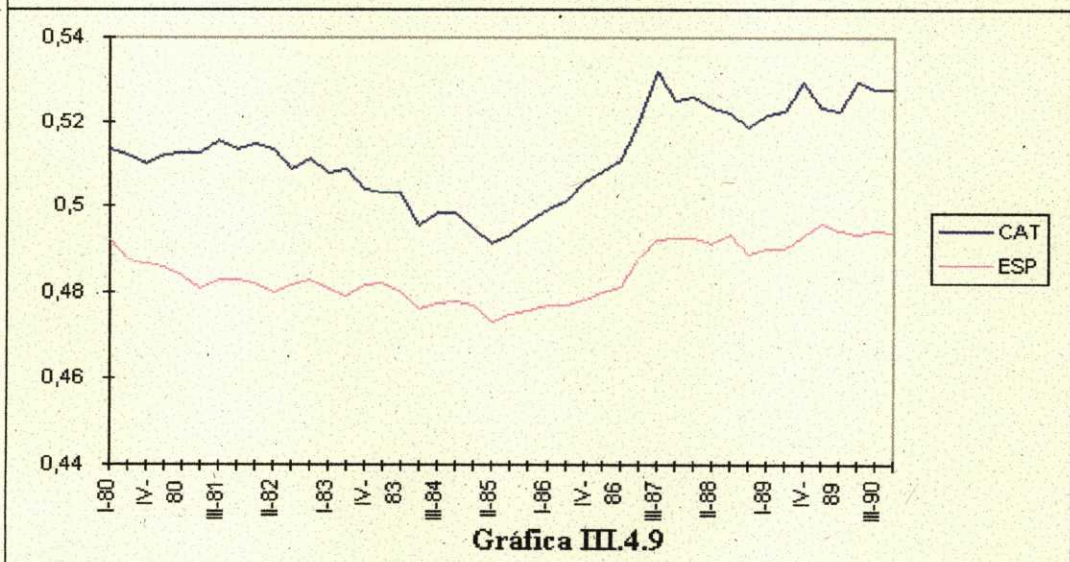
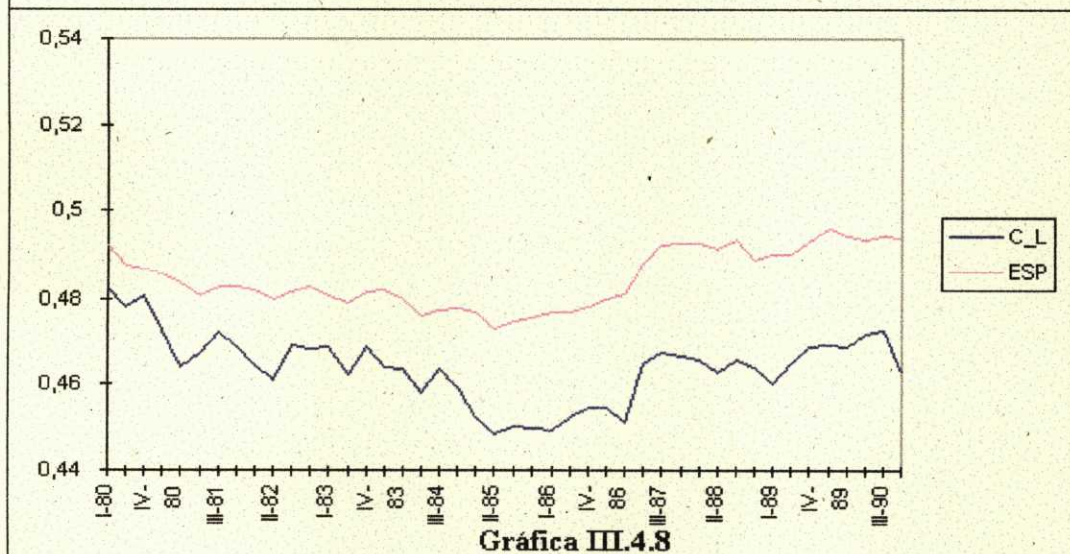
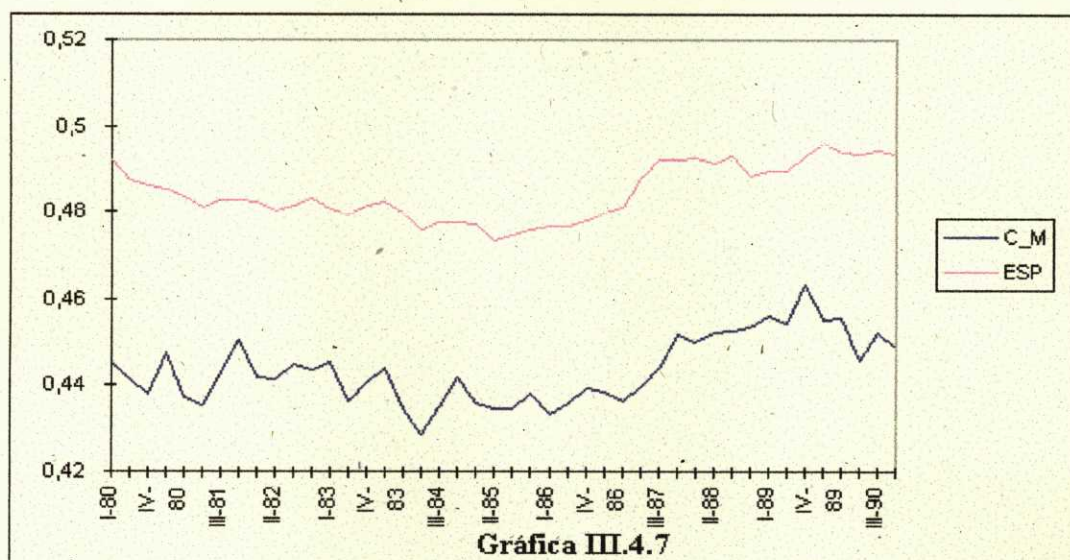
**Tasa de Actividad.
(Andalucía, Aragón y Asturias)**



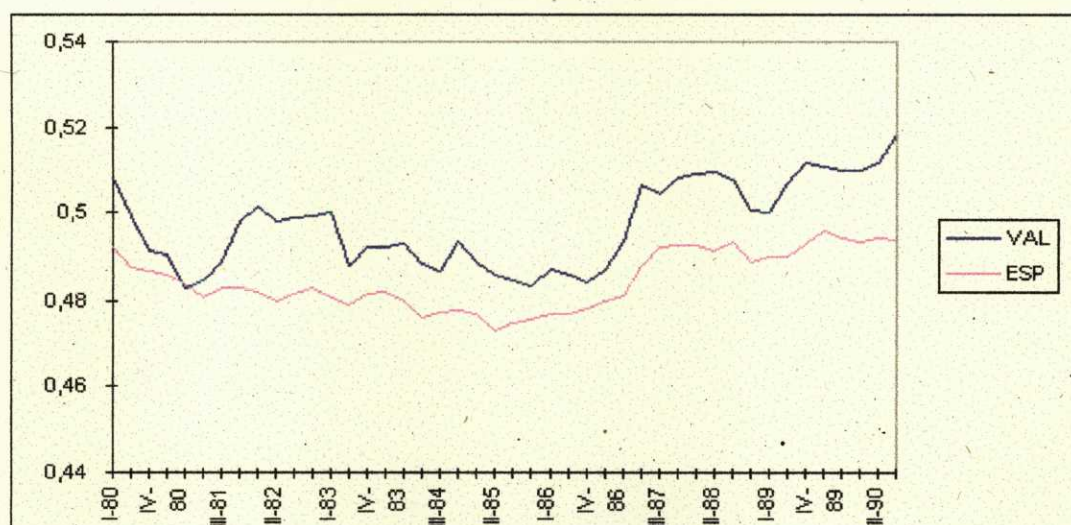
Tasa de Actividad.
(Baleares, Canarias y Cantabria)



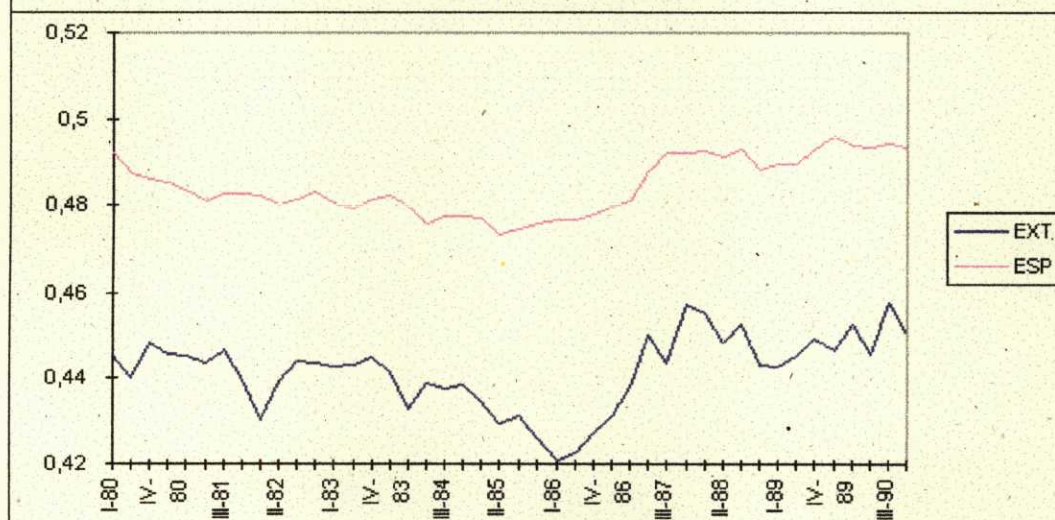
Tasa de Actividad.
(Castilla - La Mancha, Castilla y León, y Cataluña)



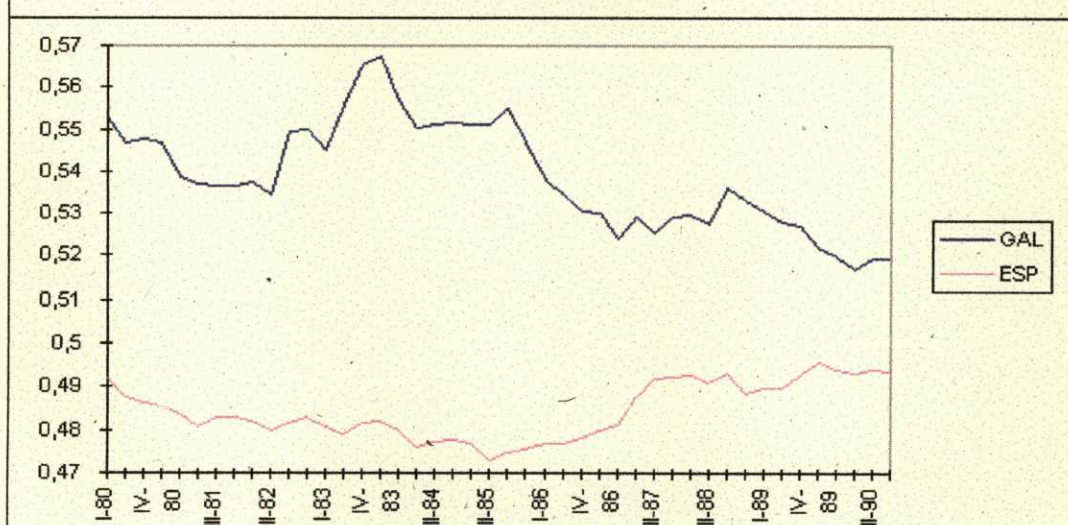
Tasa de Actividad. (Comunidad Valenciana, Extremadura y Galicia)



Gráfica III.4.10

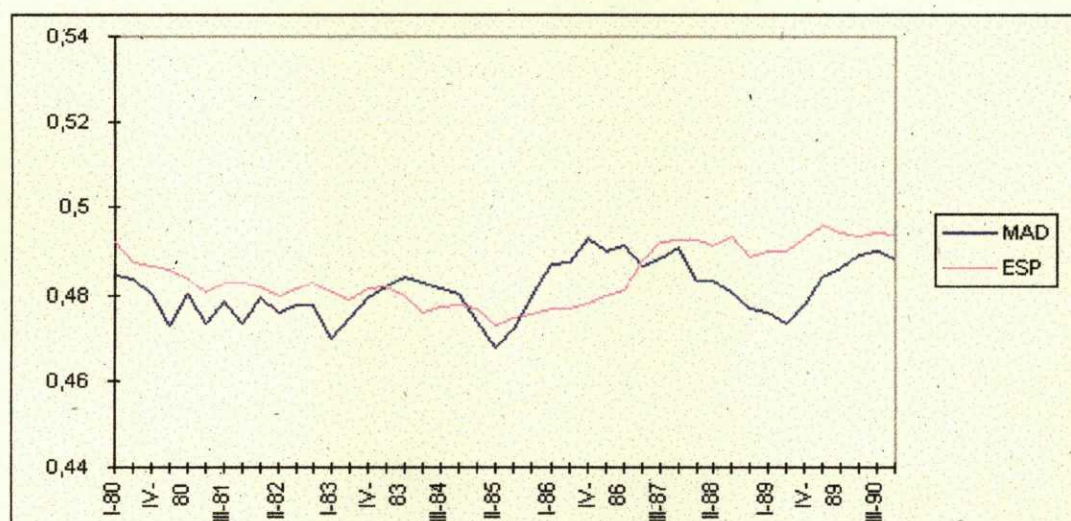


Gráfica III.4.11

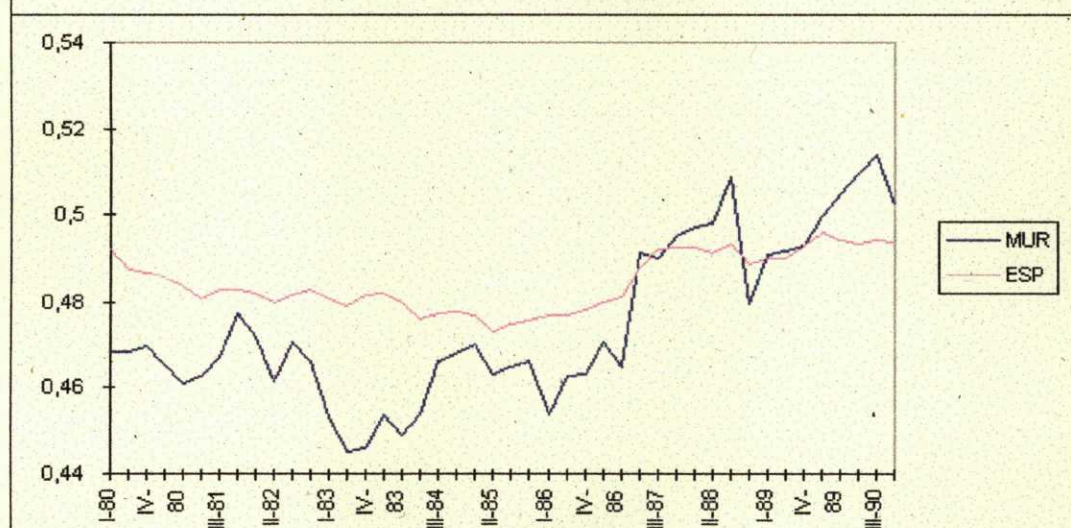


Gráfica III.4.12

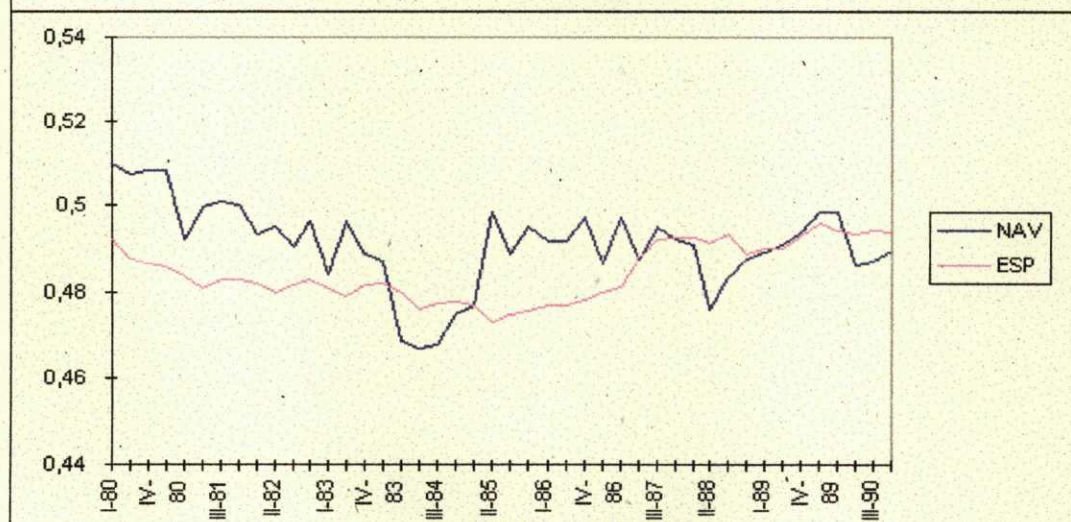
Tasa de Actividad. (Madrid, Murcia y Navarra)



Gráfica III.4.13

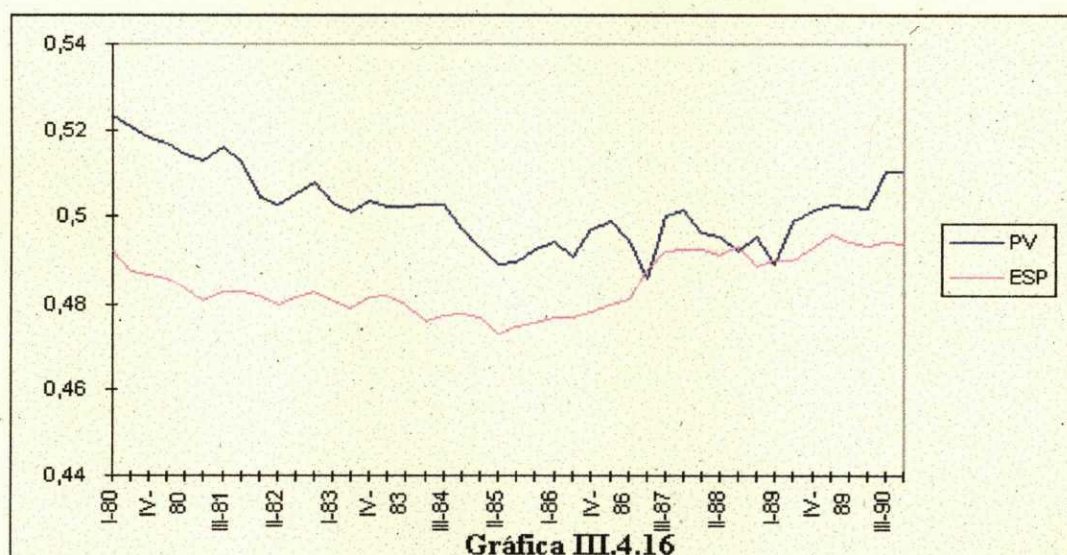


Gráfica III.4.14

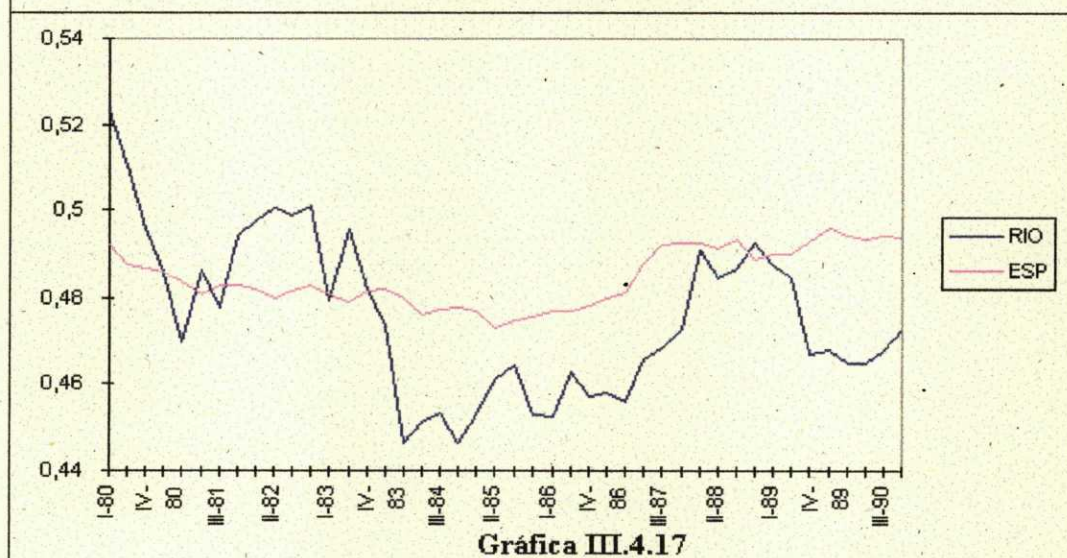


Gráfica III.4.15

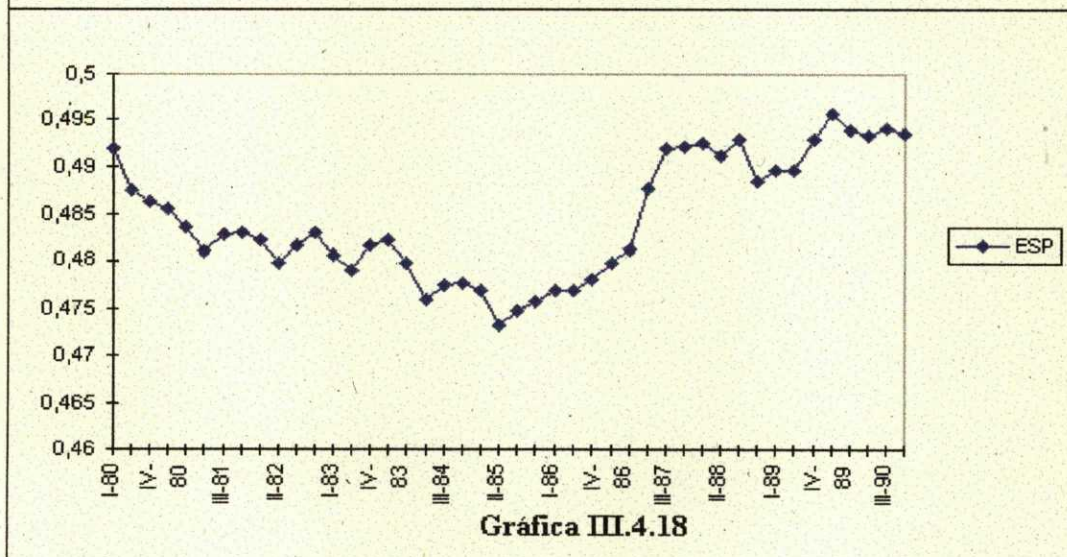
Tasa de Actividad.
(País Vasco, La Rioja y datos globales)



Gráfica III.4.16



Gráfica III.4.17



Gráfica III.4.18

APÉNDICE III.4 Contraste de los modelos I(1).

En este apéndice se incluyen las salidas de ordenador de los contrastes comentados en el *Apartado 3.2.3* de este trabajo.

El texto de este apéndice está incluido en los ficheros APIII4.doc (formato Microsoft Word 6.0) y APIII4.wp (formato Word Perfect 5.1).

APÉNDICE III.5 Resultados de las series de la oferta de trabajo.

En este apéndice se incluyen las salidas de ordenador de los análisis de las series autonómicas de la oferta de trabajo realizados en el *Apartado 3.4*.

El texto de este apéndice está incluido en los ficheros APIII5.doc (formato Microsoft Word 6.0) y APIII5.wp (formato Word Perfect 5.1).

APÉNDICE IV.1: Contraste de especificación de Hausman.

El test de especificación de Hausman es un contraste de especificación para aplicar en una generalidad de situaciones.

La idea fundamental que se propone en Hausman (1978) es que bajo la hipótesis nula, en la que se supone que el modelo fue bien especificado, existe un estimador (β_0) consistente, asintóticamente normal y asintóticamente eficiente (en el sentido de que alcanza asintóticamente la cota de Cramer-Rao) y que sin embargo deja de ser consistente bajo la hipótesis alternativa, cuando la especificación del modelo no es correcta, en cuyo caso será sesgado e inconsistente. Por otro lado se supone que se puede encontrar otro estimador (β_1) que no sea afectado por la mala especificación, y que por lo tanto será asintóticamente normal y consistente tanto bajo H_0 como bajo H_1 , si bien no será asintóticamente eficiente.

Considerando como se comportan los dos estimadores bajo las hipótesis nula y alternativa, la diferencia

$$\hat{q} = \beta_1 - \beta_0$$

en el límite diferirá de cero cuando es correcta la hipótesis alternativa, cosa que no ocurrirá si la correcta es la hipótesis nula.

Partiendo de esta idea intuitiva se puede establecer el método de contrastes de Hausman más formalmente.

Se quiere contrastar:

H_0 : el modelo fue bien especificado

con respecto a:

H_1 : la especificación del modelo no es correcta.

Para hacer uso del estadístico \hat{q} es necesario conocer su ley de distribución.

La varianza de la distribución asintótica de $\sqrt{T}\hat{q}$, $\hat{V}(\hat{q})$, precisa para el desarrollo del contraste, se demuestra en Hausman (1978) que puede ser calculada de la forma siguiente:

$$\hat{V}(\hat{q}) = \hat{V}(\beta_1 - \beta_0) = \hat{V}(\beta_1) - \hat{V}(\beta_0)$$

además, como β_0 alcanza asintóticamente la cota de Cramer-Rao, demuestra Hausman que:

$$\hat{V}(\hat{q}) = \hat{V}(\beta_1) - \hat{V}(\beta_0) \geq 0$$

en el sentido de ser una matriz definida no negativa.

Con todo lo anterior se puede elaborar un contraste general de especificación, considerando el estadístico:

$$m = T \cdot \hat{q}' \hat{V}(\hat{q})^{-1} \hat{q}$$

donde T es el tamaño de la muestra y $\hat{V}(\hat{q})$ un estimador consistente de la varianza de \hat{q} . Bajo H_0 se demuestra que este estadístico se distribuye asintóticamente como una χ^2_k , siendo k el número de parámetros que se estimaron bajo H_0 .

Por consiguiente, la regla de decisiones es directa, una vez fijado el nivel de significación

APÉNDICE IV.2 Resultados de las estimaciones mínimo-cuadráticas de los modelos.

En este apéndice se incluyen las salidas de ordenador de las regresiones mínimo-cuadráticas comentadas en el *Apartado 4.1*.

El texto de este apéndice está incluido en los ficheros APIV2.doc (formato Microsoft Word 6.0) y APIV2.wp (formato Word Perfect 5.1).

APÉNDICE IV.3 Diseño del estimador semiparamétrico: Cálculos y diagrama de flujo

En la primera parte de este apéndice se calculan los elementos de las matrices C y D obtenidos para la función δ que se ha elegido. Las otras dos partes están dedicadas al programa que realiza la estimación semiparamétrica de los modelos, presentándose en primer lugar su diagrama de flujo, y finalmente el código del programa elaborado en Turbo-Pascal.

Cálculos para la obtención del estimador semiparamétrico:

$$\left. \begin{matrix} r=1,2,\dots,M \\ i=1,2,\dots,p \\ j=1,2,\dots,q \end{matrix} \right\} \text{ ó } \left\{ \begin{matrix} i=1,2,\dots,q \\ j=q+1,2,\dots,p \end{matrix} \right.$$

$i \neq j$

$$\bar{V}_{ij}^r = \frac{\sum_{t=1}^T \sum_{m=1}^M \int x^i x^j I_r(x) \delta(x, X_{tm}) dx}{TM} =$$

$$= \frac{\sum_{t=1}^T \sum_{m=1}^M \left(\int_{X_{tm}^i - h_1}^{X_{tm}^i + h_1} \frac{I_r(x)}{2h_1} dx^1 \right) \dots \left(\int_{X_{tm}^i - h_i}^{X_{tm}^i + h_i} x^i \frac{I_r(x)}{2h_i} dx^i \right) \dots \left(\int_{X_{tm}^j - h_j}^{X_{tm}^j + h_j} x^j \frac{I_r(x)}{2h_j} dx^j \right) \dots \left(\int_{X_{tm}^p - h_p}^{X_{tm}^p + h_p} \frac{I_r(x)}{2h_p} dx^p \right)}{TM}$$

Como $I_r(x)=0$ siempre que x no esté en la r -ésima submuestra tendré:

$$\bar{V}_{ij}^r = \frac{\sum_{t=1}^T \left(\int_{X_{tr}^i - h_1}^{X_{tr}^i + h_1} \frac{dx^1}{2h_1} \right) \dots \left(\int_{X_{tr}^i - h_i}^{X_{tr}^i + h_i} x^i \frac{dx^i}{2h_i} \right) \dots \left(\int_{X_{tr}^j - h_j}^{X_{tr}^j + h_j} x^j \frac{dx^j}{2h_j} \right) \dots \left(\int_{X_{tr}^p - h_p}^{X_{tr}^p + h_p} \frac{dx^p}{2h_p} \right)}{TM} =$$

$$\begin{aligned}
 & \frac{\sum_{i=1}^T \left(\frac{2h_1}{2h_1} \right) \dots \frac{1}{2h_i} \frac{(x^i)^2}{2} \Bigg|_{X_{tr}^i - h_i}^{X_{tr}^i + h_i} \dots \frac{1}{2h_j} \frac{(x^j)^2}{2} \Bigg|_{X_{tr}^j - h_j}^{X_{tr}^j + h_j} \dots \left(\frac{2h_p}{2h_p} \right)}{TM} \\
 &= \frac{\sum_{i=1}^T \left[\frac{\left((X_{tr}^i)^2 + 2X_{tr}^i h_i + h_i^2 \right) - \left((X_{tr}^i)^2 - 2X_{tr}^i h_i + h_i^2 \right)}{4h_i} \right] \left[\frac{\left((X_{tr}^j)^2 + 2X_{tr}^j h_j + h_j^2 \right) - \left((X_{tr}^j)^2 - 2X_{tr}^j h_j + h_j^2 \right)}{4h_j} \right]}{TM} \\
 &= \frac{\sum_{i=1}^T \frac{4X_{tr}^i h_i}{4h_i} \cdot \frac{4X_{tr}^j h_j}{4h_j}}{TM} = \frac{\sum_{i=1}^T X_{tr}^i \cdot X_{tr}^j}{TM}
 \end{aligned}$$

$$\begin{aligned}
 & r=1,2,\dots,M \\
 & i=1,2,\dots,p \quad \text{ó} \quad \begin{cases} i=1,2,\dots,q \\ j=q+1,2,\dots,p \end{cases} \\
 & j=1,2,\dots,q
 \end{aligned}$$

$$i=j$$

$$\begin{aligned}
 \bar{V}_{ii}^r &= \frac{\sum_{i=1}^T \sum_{m=1}^M \int (x^i)^2 I_r(x) \delta(x, X_{im}) dx}{TM} \\
 &= \frac{\sum_{i=1}^T \sum_{m=1}^M \left(\int_{X_{im}^i - h_1}^{X_{im}^i + h_1} \frac{I_r(x)}{2h_1} dx \right) \dots \left(\int_{X_{im}^i - h_i}^{X_{im}^i + h_i} \frac{(x^i)^2}{2h_i} dx \right) \dots \left(\int_{X_{im}^i - h_p}^{X_{im}^i + h_p} \frac{I_r(x)}{2h_p} dx \right)}{TM}
 \end{aligned}$$

Como $I_r(x)=0$ siempre que x no esté en la r -ésima submuestra tendré:

$$\begin{aligned}
 \bar{V}_{ii}^r &= \frac{\sum_{i=1}^T \left(\int_{X_{tr}^i - h_1}^{X_{tr}^i + h_1} \frac{dx^1}{2h_1} \right) \dots \left(\int_{X_{tr}^i - h_i}^{X_{tr}^i + h_i} \frac{(x^i)^2}{2h_i} dx^i \right) \dots \left(\int_{X_{tr}^i - h_p}^{X_{tr}^i + h_p} \frac{dx^p}{2h_p} \right)}{TM} \\
 &= \frac{\sum_{i=1}^T \left(\frac{2h_1}{2h_1} \right) \dots \frac{1}{2h_i} \frac{(x^i)^3}{3} \Bigg|_{X_{tr}^i - h_i}^{X_{tr}^i + h_i} \dots \left(\frac{2h_p}{2h_p} \right)}{TM}
 \end{aligned}$$

$$= \frac{\sum_{t=1}^T \left[\frac{\left((X_{tr}^i)^3 + 3(X_{tr}^i)^2 h_i + 3X_{tr}^i h_i^2 + h_i^3 \right) - \left((X_{tr}^i)^3 - 3(X_{tr}^i)^2 h_i + 3X_{tr}^i h_i^2 - h_i^3 \right)}{6h_i} \right]}{TM} =$$

$$= \frac{\sum_{t=1}^T \frac{6(X_{tr}^i)^2 h_i + 2h_i^3}{6h_i}}{TM} = \frac{\sum_{t=1}^T (X_{tr}^i)^2 + \frac{h_i^2}{3}}{TM}$$

$$\begin{matrix} r=1,2,\dots,M \\ i=1,2,\dots,p \end{matrix} \left\{ \begin{matrix} i=1,2,\dots,q \\ j=1,2,\dots,q \end{matrix} \right\} \text{ ó } \left\{ \begin{matrix} i=1,2,\dots,q \\ j=q+1,2,\dots,p \end{matrix} \right\}$$

$$i=1 \text{ ó } j=1 \text{ pero } i \neq j$$

$$\bar{V}'_{ii} = \bar{V}'_{i1} = \frac{\sum_{t=1}^T \sum_{m=1}^M \int x^i I_r(x) \delta(x, X_{tm}) dx}{TM} =$$

$$= \frac{\sum_{t=1}^T \sum_{m=1}^M \left(\int_{X_{tm}^l - h_1}^{X_{tm}^l + h_1} \frac{I_r(x)}{2h_1} dx \right) \dots \left(\int_{X_{tm}^i - h_i}^{X_{tm}^i + h_i} \frac{x^i I_r(x)}{2h_i} dx \right) \dots \left(\int_{X_{tm}^p - h_p}^{X_{tm}^p + h_p} \frac{I_r(x)}{2h_p} dx \right)}{TM}$$

Como $I_r(x)=0$ siempre que x no esté en la r -ésima submuestra tendré:

$$\bar{V}'_{ii} = \bar{V}'_{i1} = \frac{\sum_{t=1}^T \left(\int_{X_{tr}^l - h_1}^{X_{tr}^l + h_1} \frac{dx^1}{2h_1} \right) \dots \left(\int_{X_{tr}^i - h_i}^{X_{tr}^i + h_i} \frac{x^i dx^i}{2h_i} \right) \dots \left(\int_{X_{tr}^p - h_p}^{X_{tr}^p + h_p} \frac{dx^p}{2h_p} \right)}{TM} =$$

$$= \frac{\sum_{t=1}^T \left(\frac{2h_1}{2h_1} \right) \dots \left[\frac{1}{2h_i} \frac{(x^i)^2}{2} \right]_{X_{tr}^i - h_i}^{X_{tr}^i + h_i} \dots \left(\frac{2h_p}{2h_p} \right)}{TM} =$$

$$= \frac{\sum_{t=1}^T \left[\frac{\left((X_{tr}^i)^2 + 2X_{tr}^i h_i + h_i^2 \right) - \left((X_{tr}^i)^2 - 2X_{tr}^i h_i + h_i^2 \right)}{4h_i} \right]}{TM} = \frac{\sum_{t=1}^T \frac{4X_{tr}^i h_i}{4h_i}}{TM} = \frac{\sum_{t=1}^T X_{tr}^i}{TM}$$

$$\begin{matrix} r=1,2,\dots,M \\ i=1,2,\dots,p \\ j=1,2,\dots,q \end{matrix} \quad \text{ó} \quad \begin{matrix} i=1,2,\dots,q \\ j=q+1,2,\dots,p \end{matrix}$$

$$i=j=1$$

$$\bar{V}'_{11} = \frac{\sum_{t=1}^T \sum_{m=1}^M \int I_r(x) \delta(x, X_{tm}) dx}{TM} =$$

$$= \frac{\sum_{t=1}^T \sum_{m=1}^M \left(\int_{X_{tm}^1 - h_1}^{X_{tm}^1 + h_1} \frac{I_r(x)}{2h_1} dx^1 \right) \dots \left(\int_{X_{tm}^i - h_i}^{X_{tm}^i + h_i} \frac{I_r(x)}{2h_i} dx^i \right) \dots \left(\int_{X_{tm}^p - h_p}^{X_{tm}^p + h_p} \frac{I_r(x)}{2h_p} dx^p \right)}{TM}$$

Como $I_r(x)=0$ siempre que x no esté en la r -ésima submuestra tendré:

$$\begin{aligned} \bar{V}'_{11} &= \frac{\sum_{t=1}^T \left(\int_{X_{tr}^1 - h_1}^{X_{tr}^1 + h_1} \frac{dx^1}{2h_1} \right) \dots \left(\int_{X_{tr}^i - h_i}^{X_{tr}^i + h_i} \frac{dx^i}{2h_i} \right) \dots \left(\int_{X_{tr}^p - h_p}^{X_{tr}^p + h_p} \frac{dx^p}{2h_p} \right)}{TM} = \\ &= \frac{\sum_{t=1}^T \left(\frac{2h_1}{2h_1} \right) \dots \left(\frac{2h_i}{2h_i} \right) \dots \left(\frac{2h_p}{2h_p} \right)}{TM} = \frac{1}{M} \end{aligned}$$

$$\begin{matrix} i=q+1,2,\dots,p \\ j=q+1,2,\dots,p \end{matrix}$$

$$i \neq j$$

$$\bar{V}_{ij} = \frac{\sum_{t=1}^T \sum_{m=1}^M \int x^i x^j \delta(x, X_{tm}) dx}{TM} =$$

$$= \frac{\sum_{t=1}^T \sum_{m=1}^M \left(\int_{X_{tm}^1 - h_1}^{X_{tm}^1 + h_1} \frac{dx^1}{2h_1} \right) \dots \left(\int_{X_{tm}^i - h_i}^{X_{tm}^i + h_i} x^i \frac{dx^i}{2h_i} \right) \dots \left(\int_{X_{tm}^j - h_j}^{X_{tm}^j + h_j} x^j \frac{dx^j}{2h_j} \right) \dots \left(\int_{X_{tm}^p - h_p}^{X_{tm}^p + h_p} \frac{dx^p}{2h_p} \right)}{TM} =$$

$$\begin{aligned}
 & \frac{\sum_{t=1}^T \sum_{m=1}^M \left(\frac{2h_1}{2h_1} \right) \dots \frac{1}{2h_i} \frac{(x^i)^2}{2} \Bigg]_{X_{tm}^i - h_i}^{X_{tm}^i + h_i} \dots \frac{1}{2h_j} \frac{(x^j)^2}{2} \Bigg]_{X_{tm}^j - h_j}^{X_{tm}^j + h_j} \dots \left(\frac{2h_p}{2h_p} \right)}{TM} \\
 &= \frac{\sum_{t=1}^T \sum_{m=1}^M \left[\left((X_{tm}^i)^2 + 2X_{tm}^i h_i + h_i^2 \right) - \left((X_{tm}^i)^2 - 2X_{tm}^i h_i + h_i^2 \right) \right]}{4h_i} \frac{\left[\left((X_{tm}^j)^2 + 2X_{tm}^j h_j + h_j^2 \right) - \left((X_{tm}^j)^2 - 2X_{tm}^j h_j + h_j^2 \right) \right]}{4h_j}}{TM} \\
 &= \frac{\sum_{t=1}^T \sum_{m=1}^M \frac{4X_{tm}^i h_i}{4h_i} \cdot \frac{4X_{tm}^j h_j}{4h_j}}{TM} = \frac{\sum_{t=1}^T \sum_{m=1}^M X_{tm}^i \cdot X_{tm}^j}{TM}
 \end{aligned}$$

$$\begin{aligned}
 i &= q+1, 2, \dots, p \\
 j &= q+1, 2, \dots, p
 \end{aligned}$$

$$i=j$$

$$\begin{aligned}
 \bar{V}_{ii} &= \frac{\sum_{t=1}^T \sum_{m=1}^M \int (x^i)^2 \delta(x, X_{tm}^i) dx}{TM} \\
 &= \frac{\sum_{t=1}^T \sum_{m=1}^M \left(\int_{X_{tm}^i - h_i}^{X_{tm}^i + h_i} \frac{dx^i}{2h_i} \right) \dots \left(\int_{X_{tm}^i - h_i}^{X_{tm}^i + h_i} (x^i)^2 \frac{dx^i}{2h_i} \right) \dots \left(\int_{X_{tm}^p - h_p}^{X_{tm}^p + h_p} \frac{dx^p}{2h_p} \right)}{TM} \\
 &= \frac{\sum_{t=1}^T \sum_{m=1}^M \left(\frac{2h_1}{2h_1} \right) \dots \frac{1}{2h_i} \frac{(x^i)^3}{3} \Bigg]_{X_{tm}^i - h_i}^{X_{tm}^i + h_i} \dots \left(\frac{2h_p}{2h_p} \right)}{TM} \\
 &= \frac{\sum_{t=1}^T \sum_{m=1}^M \left[\left((X_{tm}^i)^3 + 3(X_{tm}^i)^2 h_i + 3X_{tm}^i h_i^2 + h_i^3 \right) - \left((X_{tm}^i)^3 + 3(X_{tm}^i)^2 h_i + 3X_{tm}^i h_i^2 + h_i^3 \right) \right]}{6h_i}}{TM} \\
 &= \frac{\sum_{t=1}^T \sum_{m=1}^M \frac{6(X_{tm}^i)^2 h_i + 2h_i^3}{6h_i}}{TM} = \frac{\sum_{t=1}^T \sum_{m=1}^M \left((X_{tm}^i)^2 + \frac{h_i^2}{3} \right)}{TM}
 \end{aligned}$$

Para la matriz D tendré los siguientes elementos:

$$r=1,2,\dots,M$$

$$i=2,\dots,q$$

$$\bar{T}_i^r = \frac{\sum_{t=1}^T \sum_{m=1}^M \int Y_{tm} x^i I_r(x) \delta(x, X_{tm}) dx}{TM}$$

$$= \frac{\sum_{t=1}^T \sum_{m=1}^M Y_{tm} \left(\int_{X_{tm}^1 - h_1}^{X_{tm}^1 + h_1} \frac{I_r(x)}{2h_1} dx^1 \right) \dots \left(\int_{X_{tm}^i - h_i}^{X_{tm}^i + h_i} \frac{I_r(x)}{2h_i} dx^i \right) \dots \left(\int_{X_{tm}^p - h_p}^{X_{tm}^p + h_p} \frac{I_r(x)}{2h_p} dx^p \right)}{TM}$$

Como $I_r(x)=0$ siempre que x no esté en la r -ésima submuestra tendré:

$$\bar{T}_i^r = \frac{\sum_{t=1}^T Y_{tr} \left(\int_{X_{tr}^1 - h_1}^{X_{tr}^1 + h_1} \frac{dx^1}{2h_1} \right) \dots \left(\int_{X_{tr}^i - h_i}^{X_{tr}^i + h_i} \frac{dx^i}{2h_i} \right) \dots \left(\int_{X_{tr}^p - h_p}^{X_{tr}^p + h_p} \frac{dx^p}{2h_p} \right)}{TM}$$

$$= \frac{\sum_{t=1}^T Y_{tr} \left(\frac{2h_1}{2h_1} \right) \dots \left[\frac{1}{2h_i} \frac{(x^i)^2}{2} \right]_{X_{tr}^i - h_i}^{X_{tr}^i + h_i} \dots \left(\frac{2h_p}{2h_p} \right)}{TM}$$

$$= \frac{\sum_{t=1}^T Y_{tr} \frac{\left[\left((X_{tr}^i)^2 + 2X_{tr}^i h_i + h_i^2 \right) - \left((X_{tr}^i)^2 - 2X_{tr}^i h_i + h_i^2 \right) \right]}{4h_i}}{TM} = \frac{\sum_{t=1}^T Y_{tr} \frac{4X_{tr}^i h_i}{4h_i}}{TM} = \frac{\sum_{t=1}^T Y_{tr} X_{tr}^i}{TM}$$

$$r=1,2,\dots,M$$

$$i=1$$

$$\bar{T}_1^r = \frac{\sum_{t=1}^T \sum_{m=1}^M \int Y_{tm} I_r(x) \delta(x, X_{tm}) dx}{TM}$$

$$= \frac{\sum_{t=1}^T \sum_{m=1}^M Y_{tm} \left(\int_{X_{tm}^1 - h_1}^{X_{tm}^1 + h_1} \frac{I_r(x)}{2h_1} dx^1 \right) \dots \left(\int_{X_{tm}^i - h_i}^{X_{tm}^i + h_i} \frac{I_r(x)}{2h_i} dx^i \right) \dots \left(\int_{X_{tm}^p - h_p}^{X_{tm}^p + h_p} \frac{I_r(x)}{2h_p} dx^p \right)}{TM}$$

Como $I_r(x)=0$ siempre que x no esté en la r -ésima submuestra tendré:

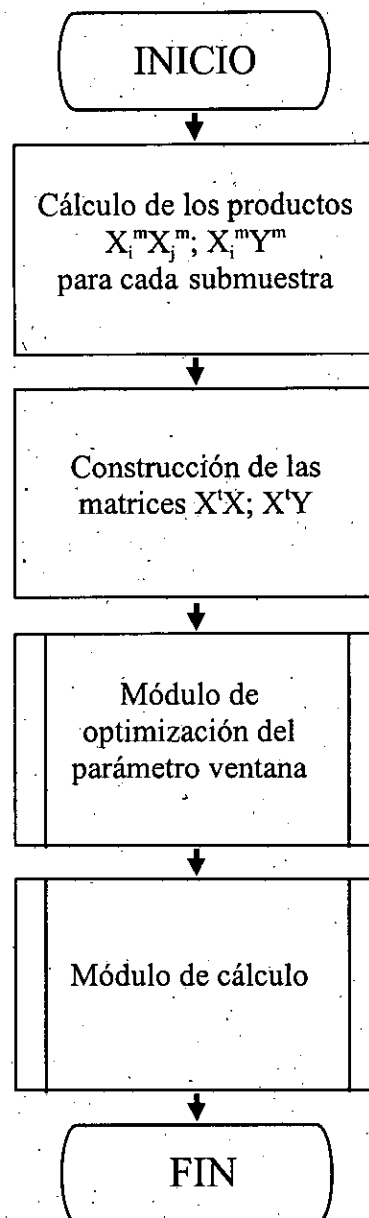
$$\begin{aligned}\bar{T}_1' &= \frac{\sum_{t=1}^T Y_{tr} \left(\int_{X_{tr}^l - h_1}^{X_{tr}^l + h_1} \frac{dx^1}{2h_1} \right) \cdots \left(\int_{X_{tr}^i - h_i}^{X_{tr}^i + h_i} \frac{dx^i}{2h_i} \right) \cdots \left(\int_{X_{tr}^p - h_p}^{X_{tr}^p + h_p} \frac{dx^p}{2h_p} \right)}{TM} = \\ &= \frac{\sum_{t=1}^T Y_{tr} \left(\frac{2h_1}{2h_1} \right) \cdots \left(\frac{2h_i}{2h_i} \right) \cdots \left(\frac{2h_p}{2h_p} \right)}{TM} = \frac{\sum_{t=1}^T Y_{tr}}{TM}\end{aligned}$$

$i=q+1, \dots, p$

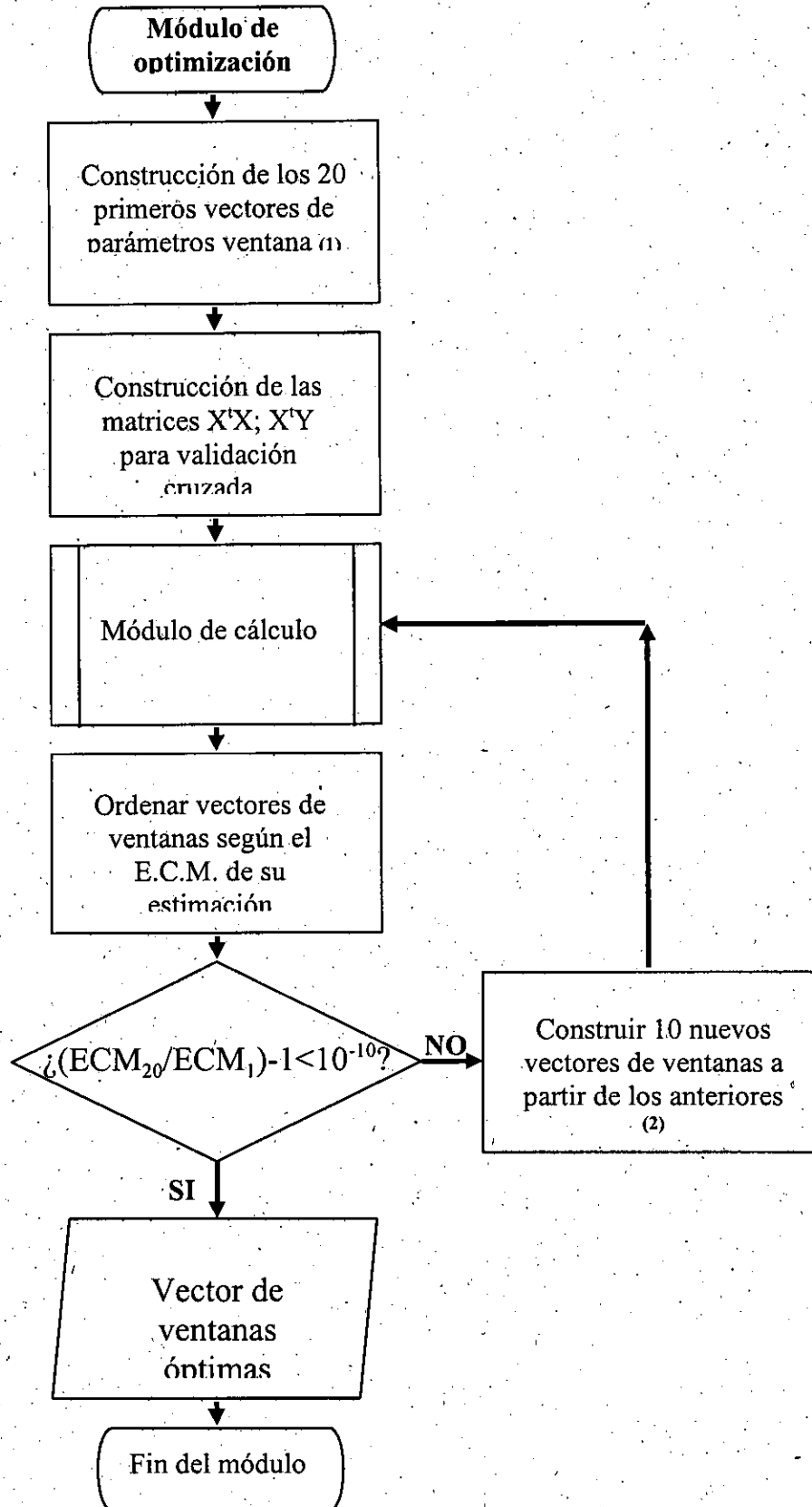
$$\begin{aligned}\bar{T}_i &= \frac{\sum_{t=1}^T \sum_{m=1}^M \int Y_{tm} x^i \delta(x, X_{tm}) dx}{TM} = \\ &= \frac{\sum_{t=1}^T \sum_{m=1}^M Y_{tm} \left(\int_{X_{tm}^l - h_1}^{X_{tm}^l + h_1} \frac{dx^1}{2h_1} \right) \cdots \left(\int_{X_{tm}^i - h_i}^{X_{tm}^i + h_i} \frac{dx^i}{2h_i} \right) \cdots \left(\int_{X_{tm}^p - h_p}^{X_{tm}^p + h_p} \frac{dx^p}{2h_p} \right)}{TM} = \\ &= \frac{\sum_{t=1}^T \sum_{m=1}^M Y_{tm} \left(\frac{2h_1}{2h_1} \right) \cdots \left[\frac{1}{2h_i} \frac{(x^i)^2}{2} \right]_{X_{tm}^i - h_i}^{X_{tm}^i + h_i} \cdots \left(\frac{2h_p}{2h_p} \right)}{TM} = \\ &= \frac{\sum_{t=1}^T \sum_{m=1}^M Y_{tm} \left[\left((X_{tm}^i)^2 + 2 X_{tm}^i h_i + h_i^2 \right) - \left((X_{tm}^i)^2 - 2 X_{tm}^i h_i + h_i^2 \right) \right]}{4 h_i TM} = \\ &= \frac{\sum_{t=1}^T \sum_{m=1}^M Y_{tm} \frac{4 X_{tm}^i h_i}{4 h_i}}{TM} = \frac{\sum_{t=1}^T \sum_{m=1}^M Y_{tm} X_{tm}^i}{TM}\end{aligned}$$

Diagrama de flujo.

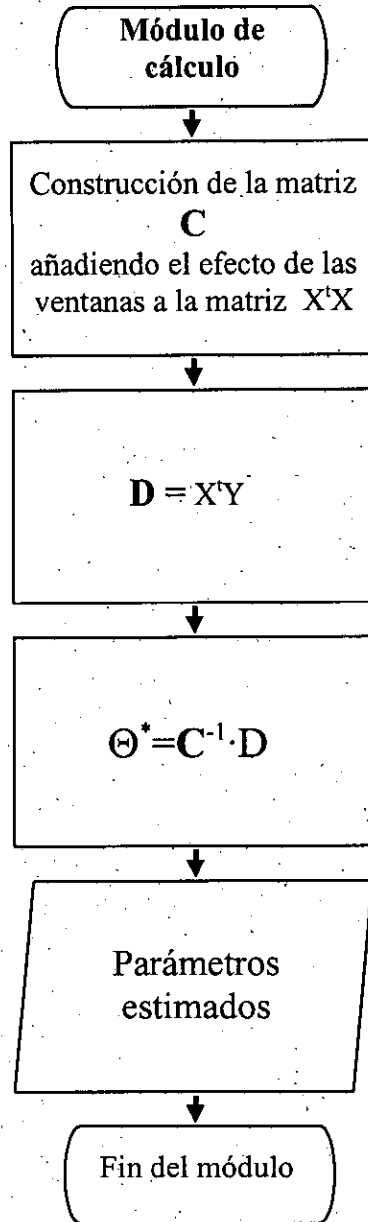
Esta parte muestra el diagrama de flujo del proceso que se sigue para la estimación de los parámetros mediante el método de estimación elaborado en este trabajo. Este proceso se ha dividido en dos partes: la primera es el proceso mediante el cual se intenta buscar el parámetro de suavización óptimo; la segunda es la estimación propiamente dicha.



MÓDULO DE OPTIMIZACIÓN DE LA VENTANA



MÓDULO DE CÁLCULO



Comentarios:

- (1) De los 20 vectores de ventanas, 18 son generados aleatoriamente, tomando valores entre 10-7 y 10; un vector tiene todas las ventanas iguales a 1.5, y otro todas las ventanas iguales a cero (es la estimación mínimo cuadrática).
- (2) Para obtener los nuevos vectores se escogen dos de los ya calculados, dándole mayor probabilidad de ser elegidos a los que tienen menor E.C.M. La coordenada i -ésima del vector combinación se elige aleatoriamente en el siguiente intervalo:

$$(\max (0, h_i - 2 \cdot (H_i - h_i)), H_i + 2 \cdot (H_i - h_i))$$

Siendo h_i la coordenada i -ésima **menor** entre las de los dos vectores elegidos,

H_i : la coordenada i -ésima **mayor** entre las de los dos vectores elegidos.

Código del programa.

Este apéndice incluye el código fuente del programa que realiza la estimación mediante suavización paramétrica previa en datos de panel. Esta realizado en TURBO PASCAL 5.5 para ordenadores PC.

Esta versión realiza la estimación para el caso en el que los parámetros no varían a través de las submuestras. Cambiando los valores que aparecen en las constantes se puede adaptar para los demás casos.

OPTIMIZACIÓN DE LA VENTANA

```
program xanela;

uses crt;

const np=3;      {numero de pmts.}
      nm=1;      {numero de submostras}
      nd=100;    {numero de datos de cada submostra}
      npc=0;
      npr=(nm-1)*npc+np;

type mtxnng=array [1..np,1..20] of real;
      mtp=array[1..npr] of real;
      mtxn=array [1..np] of real;
      mtg=array[1..npr,1..npr] of real;
      mg=array[1..np,1..nd] of real;
      medx=array[1..nm,1..np,1..np] of real;
      medxy=array[1..nm,1..np] of real;
      mtrs=array [1..20] of real;

var fmat           :file of mtp;
    fres          :file of real;
    fmat2         :file of mtg;
    k,ii,jj,i3,j3,s,i,j,d,kd,tk :integer;
    cs            :real;
    md, matd      :mtp;
    mc matc,      :mtg;
    nr,rs         :mtrs;
    ht            :mtxn;
    h,nh          :mtxnng;
    xm            :medx;
    xym           :medxy;

procedure ler_datos(var xm:medx;var xmy:medxy);
var ys:real;
    x :mg;
    fx:file of real;
```

```

begin
  {Inicializar los valores};
  for s:=1 to nm do for i:=1 to np do
    begin {i}
      xym[s,i]:=0;
      for j:=1 to np do xm[s,i,j]:=0;
    end; {i}

    {Asignación de los valores a X e Y}
    assign(fx,'datosd'); reset(fx);
    for s:=1 to nm do
      begin
        for d:=1 to nd do
          begin
            for i:=1 to np do read(fx,x[i,d]);
            for i:=1 to np do for j:=1 to np do
xm[s,i,j]:=xm[s,i,j]+x[i,d]*x[j,d];
            read(fx,ys);
            for i:=1 to np do xym[s,i]:=xym[s,i]+ys*x[i,d];
          end; {d}
        end; {s}
      close(fx);
    end; {ler}

procedure construir_c(xm:medx;var c2:mtg);
var cs,rl:real;
    i2,j2:integer;

begin
  for i2:=1 to npr do for j2:=1 to npr do c2[i2,j2]:=0;
  for i2:=1 to npc do
    begin
      for j2:=1 to npc do for s:=1 to nm do c2[npc*(s-1)+i2,npc*(s-
1)+j2]:=xm[s,i2,j2]/(nm*nd-1);
      for j2:=npc+1 to np do for s:=1 to nm do c2[npc*(s-
1)+i2,npc*(nm-1)+j2]:=xm[s,i2,j2]/(nm*nd-1);
    end; {i2}
    for i2:=npc+1 to np do
      begin
        for j2:=1 to npc do for s:=1 to nm do c2[npc*(nm-1)+i2,npc*(s-
1)+j2]:=xm[s,i2,j2]/(nm*nd-1);
        for j2:=npc+1 to np do
          begin
            rl:=0;
            for s:=1 to nm do rl:=rl+xm[s,i2,j2];
            c2[npc*(nm-1)+i2,npc*(nm-1)+j2]:=rl/(nm*nd-1);
          end; {j2}
        end; {i2}
      end; {construir_c}

procedure construir_d(xm:medx;xym:medxy;var d:mtg);
var rl:real;
    i2:integer;

begin
  for i2:=1 to npc do for s:=1 to nm do d[npc*(s-
1)+i2]:=xym[s,i2]/(nm*nd-1);
  for i2:=npc+1 to np do
    begin
      rl:=0;

```

```

    for s:=1 to nm do r1:=r1+xym[s,i2]/(nm*nd-1);
    d[npc*(nm-1)+i2]:=r1;
  end; {i2}
end; {construir_d}

procedure matrices(var mtc:mtg;var mtd:mtp;cs2:real);
var x      :mg;
    fx     :file of mg;
    md2    :mtp;
    mc2    :mtg;
    fy     :file of real;
    id,jd:integer;
    vi     :real;

begin
  assign(fx,'datosd');assign(fy,'datosy');
  reset(fx);reset(fy);
  seek(fx,i3-1);seek(fy,(i3-1)*nd+j3-1);
  read(fx,x);
  read(fy,vi);
  for id:=1 to npc do for jd:=1 to npc do mc2[npc*(i3-1)+id,npc*(i3-1)+jd]:=mtc[npc*(i3-1)+id,npc*(i3-1)+jd]-(x[id,j3]*x[jd,j3]/cs2);
  for id:=1 to npc do for jd:=npc+1 to np do mc2[npc*(i3-1)+id,npc*(nm-1)+jd]:=mtc[npc*(i3-1)+id,npc*(nm-1)+jd]-(x[id,j3]*x[jd,j3]/cs2);
  for id:=npc+1 to np do for jd:=1 to npc do mc2[npc*(nm-1)+id,npc*(i3-1)+jd]:=mtc[npc*(nm-1)+id,npc*(i3-1)+jd]-(x[id,j3]*x[jd,j3]/cs2);
  for id:=npc+1 to np do for jd:=npc+1 to np do mc2[npc*(nm-1)+id,npc*(nm-1)+jd]:=mtc[npc*(nm-1)+id,npc*(nm-1)+jd]-(x[id,j3]*x[jd,j3]/cs2);
  for id:=1 to npc do md2[npc*(i3-1)+id]:=mtd[npc*(i3-1)+id]-(x[id,j3]*vi/cs2);
  for id:=npc+1 to np do md2[npc*(nm-1)+id]:=mtd[npc*(nm-1)+id]-(x[id,j3]*vi/cs2);
  close(fx);close(fy);
  mtd:=md2;mtc:=mc2;
end; {matrices}

procedure cargar(h2:mtx;var c2:mtg);
var aux,i2,j2:integer;

begin
  for i2:=0 to nm-1 do for j2:=2 to npc do
    c2[i2*npc+j2,i2*npc+j2]:=c2[i2*npc+j2,i2*npc+j2]+nd*h2[j2-1]*h2[j2-1]/3;
    i2:=nm*npc+1;
    aux:=npc;
    if npc>0 then c2[i2,i2]:=c2[i2,i2]+nm*nd*h2[aux]*h2[aux]/3;
    for i2:=nm*npc+2 to npr do c2[i2,i2]:=c2[i2,i2]+nm*nd*h2[i2-1-(nm-1)*npc]*h2[i2-1-(nm-1)*npc]/3;
  end; {cargar}

procedure cambio(var a,b:real);
var aux:real;

begin
  aux:=a; a:=b; b:=aux;
end; {cambio}

```

```

procedure inversa(var m:mtg);
{invierte una matriz de dimensión npr x npr}
var aux      :mtg;
    a,b,c,j2,i2 :integer;
    pvt,cmb     :real;
    cl,c2       :array [1..npr] of integer;

begin
    for i2:=1 to npr do {construye la matriz auxiliar}
        for j2:=1 to npr do if i2=j2 then aux[i2,j2]:=1 else
aux[i2,j2]:=0;
        for s:=1 to npr do
            begin
                cl[s]:=0;
                c2[s]:=0;
            end;
        s:=0;
        for a:=1 to npr do
            begin
                if m[a,a]=0 then
                    begin
                        d:=a+1;
                        while (m[a,d]=0) and (d<=npr) do d:=d+1;
                        if d>npr then writeln('non tén inversa')
                        else
                            begin
                                cambio(m[c,a],m[c,d]);
                                s:=s+1;cl[s]:=a;c2[s]:=d;
                            end;
                        end; {if de m[a,a]=0}
                if m[a,a]<>1 then
                    begin
                        pvt:=m[a,a];
                        for b:=1 to npr do m[a,b]:=m[a,b]/pvt;
                        for b:=1 to npr do aux[a,b]:=aux[a,b]/pvt;
                    end; {if}
                for b:=1 to npr do
                    if b<>a then
                        if m[b,a]<>0 then
                            begin
                                pvt:=m[b,a];
                                for c:=1 to npr do m[b,c]:=m[b,c]-pvt*m[a,c];
                                for c:=1 to npr do aux[b,c]:=aux[b,c]-pvt*aux[a,c];
                            end; {if}
                        end; {a}
                for i2:=s downto 1 do
                    for d:=1 to npr do cambio(aux[cl[i2],d],aux[c2[i2],d]);
                for i2:=1 to npr do for j2:=1 to npr do m[i2,j2]:=aux[i2,j2];
            end; {inversa}

procedure multgp(m1:mtg;m2:mtp;var m3:mtp);
var a,c:integer;

begin
    for a:=1 to npr do
        begin
            m3[a]:=0;
            for c:=1 to npr do m3[a]:=m3[a]+m1[a,c]*m2[c];
        end; {a}
    end; {multgp}

```

```

procedure resolver(c:mtg;d:mtp;var tt:mtp);
var maux :mtg;
    ij,ji:integer;
    k      :mtxn;

```

```

begin
    maux:=c;
    inversa(maux);
    multgp(maux,d,tt);
end; {resolver}

```

```

procedure diferencia(var dif:real;tt:mtp);
{Calcula el ECM para unos parámetros estimados}

```

```

var id,is,iv :integer;
    x          :mg;
    fx         :file of mg;
    fy         :file of real;
    sum,tot,vi :real;

```

```

begin
    assign(fx,'datos');assign(fy,'datosy');
    reset(fx);reset(fy);
    dif:=0;
    for is:=1 to nm do
        begin
            read(fx,x);
            for iv:=1 to nd do
                begin
                    sum:=0;
                    read(fy,vi);
                    for id:=1 to npc do sum:=sum+x[id,iv]*tt[npc*(is-1)+id];
                    for id:=npc+1 to np do sum:=sum+x[id,iv]*tt[npc*(nm-1)+id];
                    tot:=vi-sum;
                    dif:=dif+tot*tot;
                end; {iv}
            end; {is}
        close(fx);close(fy);
    end; {diferencia}

```

```

procedure operar(mc2:mtg;md2:mtp;ht:mtxn;var rt:real);
var tt :mtp;
    dif:real;

```

```

begin
    cargar(ht,mc);
    resolver(mc2,md2,tt);
    diferencia(dif,tt);
    rt:=rt+dif*dif;
end; {operar}

```

```

procedure ordenar(var hh:mtxng;var rr:mtrs);
var csf,cdf,cs,cd,cnt,jk:integer;

```

```

begin
    csf:=19;cdf:=2;cnt:=1;cs:=cnt-1;
    while cdf<csf do
        begin
            while cnt<=csf do

```



```

begin
  if rr[cnt]>rr[cnt+1] then
    begin
      cs:=cnt;
      cambio(rr[cnt],rr[cnt+1]);
      for jk:=1 to np-1 do cambio(hh[jk,cnt],hh[jk,cnt+1]);
    end; {if}
    cnt:=cnt+1;
  end; {while cnt}
  csf:=cs;cnt:=cs;cd:=csf;
  if cdf<csf then
    begin
      while cdf<=cnt do
        begin
          if rr[cnt-1]>rr[cnt] then
            begin
              cd:=cnt;
              cambio(rr[cnt],rr[cnt-1]);
              for jk:=1 to np-1 do cambio(hh[jk,cnt],hh[jk,cnt-1]);
            end; {if}
            cnt:=cnt-1;
          end; {while cnt}
          cdf:=cd;cnt:=cd;cs:=cdf;
        end {if}
      end; {while}
    end; {ordenar}

procedure valor(var aa:real)
begin
  case random(6) of
    0:aa:=random(1000)*1e-2;
    1:aa:=random(1000)*1e-3;
    2:aa:=random(1000)*1e-4;
    3:aa:=random(1000)*1e-5;
    4:aa:=random(1000)*1e-6;
    5:aa:=random(1000)*1e-7;
  end;
end; {valor}

procedure novo_h(nota:integer;var hh2:mtxng);
type matint=array[1..20] of integer;
var w, u:integer;
    q :matint;

begin;
  {Determina los pares de vectores de ventanas que generan el nuevo vector}
  for u:=1 to 4 do q[u]:=1;
  for u:=5 to 6 do q[u]:=2;
  for u:=7 to 8 do q[u]:=random(2)+3;
  q[9]:=random(4)+6;
  q[10]:=random(9)+11;
  q[11]:=2;
  q[12]:=random(2)+3;
  q[13]:=random(4)+6;
  q[14]:=random(9)+11;
  q[15]:=random(6)+3;
  q[16]:=random(10)+10;
  q[17]:=random(6)+3;
  q[18]:=random(10)+10;
  q[19]:=random(14)+6;

```

```

q[20]:=random(19)+1;

{Realiza las combinaciones aleatorias de vectores}
for w:=1 to 10 do for u:=1 to np-1 do
begin
  if h[u,q[w]]<h[u,q[10+w]] then
    if 3*h[u,q[w]]>2*h[u,q[10+w]] then hh2[u,w]:= 3*h[u,q[w]]-
2*h[u,q[w+10]]+random*5*(h[u,q[w+10]]-h[u,q[w]])
    else hh2[u,w]:=random*(3*h[u,q[w+10]]-2*h[u,q[w]])
  else
    if 3*h[u,q[w+10]]>2*h[u,q[w]] then hh2[u,w]:=
3*h[u,q[w+10]]-2*h[u,q[w]]+random*5*(h[u,q[w]]-h[u,q[w+10]])
    else hh2[u,w]:=random*(3*h[u,q[w]]-2*h[u,q[w+10]]);
end; {u}
end; {novo_h}

```

```

procedure buscar(var a,b:integer;rs2:mtrs;rr2:real;var
cnt2:integer);
var nmr:integer;
begin
  nmr:=(a+b) div 2;
  if rs2[nmr]=rr2 then cnt2:=nmr
  else
    begin
      if rs2[nmr]>rr2 then b:=nmr-1
      else a:=nmr+1;
      if b>a then buscar(a,b,rs2,rr2,cnt2)
      else if rs2[a]>rr2 then cnt2:=a
      else cnt2:=a+1;
    end; {else}
  end; {buscar}

```

```

procedure situar(hh:mtx;rr:real);
var cnt,ei,es,i2,j2:integer;

begin
  cnt:=0;ei:=1;es:=19;
  buscar(ei,es,rs,rr,cnt);
  for i2:=20 downto cnt+1 do
    begin
      rs[i2]:=rs[i2-1];
      for j2:=1 to np-1 do h[j2,i2]:=h[j2,i2-1];
    end;
  rs[cnt]:=rr;
  for j2:=1 to np-1 do h[j2,cnt]:=hh[j2];
end; {situar}

```

```

begin
  randomize;
  ler_datos(xm,xym);

  {Inicializar valores y ficheros}
  cs:=nm*nd-1;
  for i:=1 to 20 do rs[i]:=0;
  assign(fmat,'mataux'); assign(fmat2,'mataux2');
  rewrite(fmat); rewrite(fmat2);

  {Construye las matrices  $X^tX$  y  $X^tY$ }
  construir_c(xm,matc); construir_d(xm,xym,matd);

```

```

{Genera valores aleatorios para 18 vectores de ventanas}
for i:=1 to 18 do for j:=1 to np-1 do valor(h[j,i]);

{Completa la matriz de ventanas}
for i:=1 to np-1 do h[i,19]:=1.5;
for i:=1 to np-1 do h[i,20]:=0;

{El siguiente bucle realiza nd*nm estimaciones de los parámetros
utilizando grupos de datos de los que se suprime uno diferente cada
vez}
for i3:=1 to nm do for j3:=1 to nd do
begin
mc:=matc;
md:=matd;

{Construye y guarda en ficheros matrices  $X^tX$  y  $X^tY$  en las que no
se considera uno de los datos}
matrices(mc,md,cs);
write(fmat,md);write(fmat2,mc);

{El bucle estima los parámetros usando los 20 vectores de ventanas,
guarda los correspondientes valores del E.C.M. estimado en la matriz z}
for k:=1 to 20 do
begin
for j:=1 to np-1 do ht[j]:=h[j,k];
operar(mc,md,ht,rs[k]);
end;
end; {j3,i3}
close(fmat);close(fmat2);

{Se ordenan la matriz de ventanas en función de sus E.C.M.}
ordenar(h,rs);

{Se repite el proceso de generar ventanas, estimar y ordenarlas según su
E.C.M. entrando en la matriz z si este valor está entre los 20 menores}
for kd:=1 to 5 do
begin

{Generación aleatoria de ventanas}
for i:=1 to 10 do for j:=1 to np-1 do valor(nh[j,i]);
for tk:=1 to 20 do nr[tk]:=0;
for tk:=1 to 10 do
begin
for j:=1 to np-1 do ht[j]:=nh[j,tk];
reset(fmat);reset(fmat2);

(nm*nd estimaciones de parámetros con las nuevas ventanas)
for ii:=1 to nm do for jj:=1 to nd do
begin
read(fmat,md);read(fmat2,mc);
operar(mc,md,ht,nr[tk]);
end;
close(fmat);close(fmat2);

{Si el ECM está entre los 20 mejores entra la n-pla de
ventanas entra en la matriz z}
if nr[tk]<rs[20] then situar(ht,nr[tk]);
end; {tk}
end; {kd}

{A partir de ahora se repite el proceso anterior, pero en lugar de
obtener valores aleatoriamente se combinan entre si las veinte ventanas}

```

óptimas para generar nuevas combinaciones. Se detiene el proceso cuando la diferencia entre el mejor ECM y el vigésimo es inferior a 10^{-10}

```
while ((rs[20]/rs[1])-1>1e-10) do
begin
  {Aquí se combinan los valores de las mejores ventanas, para
  obtener nuevas n-plas}
  novo_h(nh);
  for tk:=1 to 20 do nr[tk]:=0;
  for tk:=1 to 10 do
  begin
    for j:=1 to np-1 do ht[j]:=nh[j,tk];
    reset(fmat);reset(fmat2);
    for ii:=1 to nm do for jj:=1 to nd do
    begin
      read(fmat,md);read(fmat2,mc);
      operar(mc,md,ht,nr[tk]);
    end;
    close(fmat);close(fmat2);
    if nr[tk]<rs[20] then situar(ht,nr[tk]);
  end; {tk}
end; {while}

{Guarda en el fichero 'xanelas' los parámetros de suavización obtenidos}
assign(fres,'xanelas');
rewrite(fres);
for k:=1 to np-1 do write(fres,h[k,1]);
close(fres);
end.
```

ESTIMACIÓN SEMIPARAMÉTRICA

```
program noia; {9-IV-96}

uses crt;

const np=4;      {numero de pmts.}
      nm=17;     {numero de submostras}
      nd=43;     {numero de datos de cada submostra}
      npc=0;     {numero de parámetros que cambian en cada submostra}
      npr=(nm-1)*npc+np; {numero total de parámetros diferentes a
      estimar}

type mxan=array[1..np-1] of real;
      mtp=array[0..npr] of real;
      mg=array[1..np,1..nd] of real;
      mtg=array[1..npr,1..npr] of real;
      medx=array[1..nm,1..np,1..np] of real;
      medxy=array[1..nm,1..np] of real;

var s,d,i,j :integer;
    h       :mxan;
    tt,matd :mtp;
    f,f2    :file of real;
    matc    :mtg;
    xm      :medx;
    xym     :medxy;
```

```

procedure ler_datos(var xm:medx;var xym:medxy);
var ys:real;
    x :mg;

begin
  {Inicializar los valores};
  for s:=1 to nm do for i:=1 to np do
    begin {i}
      xym[s,i]:=0;
      for j:=1 to np do xm[s,i,j]:=0;
    end; {i}
  {Asignación de los valores a X e Y}
  assign(f2,'datos');
  reset(f2);
  for s:=1 to nm do
    begin
      for d:=1 to nd do
        begin
          x[1,d]:=1;
          for i:=2 to np do read (f2, x[i,d]);
          for i:=1 to np do for j:=1 to np do xm[s,i,j]:=
xm[s,i,j]+x[i,d]*x[j,d];
          read(f2,ys);
          for i:=1 to np do xym[s,i]:=xym[s,i]+ys*x[i,d];
        end; {d}
      end; {s}
    close(f2);
  end; {xerar}

```

```

procedure ler_xanelas(var h:mxan);
{lee los parámetros de suavización optimizados}
var fh:file of real;

begin
  assign(fh,'xanelas'); reset(fh);
  for s:=1 to np-1 do read (fh,h[s]);
  close(fh);
end; {ler_xanelas}

```

```

procedure construir_d(xm:medx;xym:medxy;var d:mtg);
var rl:real;

begin
  for i:=1 to npc do for s:=1 to nm do d[npc*(s-
1)+i]:=xym[s,i]/(nm*nd);
  for i:=npc+1 to np do
    begin
      rl:=0;
      for s:=1 to nm do rl:=rl+xym[s,i]/(nm*nd);
      d[npc*(nm-1)+i]:=rl;
    end; {i}
  end; {construir_d}

```

```

procedure construir_c(xm:medx;var c:mtg);
var rl :real;
    aux:integer;

begin
  for i:=1 to npr do for j:=1 to npr do c[i,j]:=0;

```

```

for i:=1 to npc do
begin
  for j:=1 to npc do for s:=1 to nm do c[npc*(s-1)+i,npc*(s-
1)+j]:=xm[s,i,j]/(nm*nd);
  for j:=npc+1 to np do for s:=1 to nm do c[npc*(s-1)+i,npc*(nm-
1)+j]:=xm[s,i,j]/(nm*nd);
  end; {i}
  for i:=npc+1 to np do
  begin
    for j:=1 to npc do for s:=1 to nm do c[npc*(nm-1)+i,npc*(s-
1)+j]:=xm[s,i,j]/(nm*nd);
    for j:=npc+1 to np do
    begin
      rl:=0;
      for s:=1 to nm do rl:=rl+xm[s,i,j];
      c[npc*(nm-1)+i,npc*(nm-1)+j]:=rl/(nm*nd);
    end; {j}
  end; {i}
  for i:=0 to nm-1 do for j:=2 to npc do c[i*npc+j,i*npc+j]:=
c[i*npc+j,i*npc+j]+nd*h[j-1]*h[j-1]/3;
  i:=nm*npc+1;
  aux:=npc;
  if npc>0 then c[i,i]:=c[i,i]+nm*nd*h[aux]*h[aux]/3;
  for i:=nm*npc+2 to npr do c[i,i]:=c[i,i]+nm*nd*h[i-1-(nm-
1)*npc]*h[i-1-(nm-1)*npc]/3;
  end; {construir_c}

```

```

procedure inversa(var m:mtg);

```

```

var aux      :mtg;
  a,b,c      :integer;
  pvt,cmb    :real;
  cl,c2      :array [1..npr] of integer;

```

```

begin
  for i:=1 to npr do {construe a matriz auxiliar}
  for j:=1 to npr do if i=j then aux[i,j]:=1 else aux[i,j]:=0;
  for s:=1 to npr do
  begin
    cl[s]:=0;
    c2[s]:=0;
  end;
  s:=0;
  for a:=1 to npr do
  begin
    if m[a,a]=0 then
    begin
      d:=a+1;
      while (m[a,d]=0) and (d<=npr) do d:=d+1;
      if d>npr then writeln('non ten inversa')
      else
      begin
        for c:=1 to npr do
        begin
          cmb:=m[c,a];
          m[c,a]:=m[c,d];
          m[c,d]:=cmb;
        end;
        s:=s+1;cl[s]:=a;c2[s]:=d;
      end;
    end; {if de m[a,a]=0}
    if m[a,a]<>1 then

```

```
begin
  pvt:=m[a,a];
  for b:=1 to npr do m[a,b]:=m[a,b]/pvt;
  for b:=1 to npr do aux[a,b]:=aux[a,b]/pvt;
end; {if}
for b:=1 to npr do
  if b<>a then
    if m[b,a]<>0 then
      begin
        pvt:=m[b,a];
        for c:=1 to npr do m[b,c]:=m[b,c]-pvt*m[a,c];
        for c:=1 to npr do aux[b,c]:=aux[b,c]-pvt*aux[a,c];
      end; {if}
    end; {a}
  for i:=s downto 1 do
    for d:=1 to npr do
      begin
        cmb:=aux[c1[i],d];
        aux[c1[i],d]:=aux[c2[i],d];
        aux[c2[i],d]:=cmb;
      end;
    for i:=1 to npr do for j:=1 to npr do m[i,j]:=aux[i,j];
  end; {inversa}

procedure multgp(m1:mtg;m2:mtp;var m3:mtp);
var a,c:integer;

begin
  for a:=1 to npr do
    begin
      m3[a]:=0;
      for c:=1 to npr do m3[a]:=m3[a]+m1[a,c]*m2[c];
    end; {a}
  end; {multgp}

procedure resolver(var c:mtg;d:mtp;var tt:mtp);

begin
  inversa(c);
  multgp(c,d,tt);
end;

begin
  ler_datos(xm,xym);
  ler_xanelas(h);
  construir_d(xm,xym,matd);
  construir_c(xm,matc);
  resolver(matc,matd,tt);
  assign(f,'result');rewrite(f);
  for i:=1 to np do write(f,tt[i]);
  close(f);
end.
```

APÉNDICE IV.4 Resultados de las estimaciones semiparamétricas de los modelos.

En este apéndice se incluyen las salidas de ordenador de las regresiones semiparamétricas y mínimo-cuadráticas comentadas en el *Apartado 4.2*.

El texto de este apéndice está incluido en los ficheros APIV4.doc (formato Microsoft Word 6.0) y APIV4.wp (formato Word Perfect 5.1).

Bibliografía:

Agüero, I.; A. Olano (1988) "Oferta de trabajo de jóvenes. Aspectos demográficos y económicos". *Economía y Sociología del trabajo. Los jóvenes y el empleo*. Ministerio de Trabajo y Seguridad Social.

Aguiló, E. (1990) "La economía balear al final de la gran expansión". *Papeles de economía española*. 45: 137-150.

Alcaide Inchausti, J. (1992) "Comportamiento económico de las comunidades autónomas españolas (1986-1991)". *Papeles de economía española*. 51: 2-33.

Alcaide Inchausti, J; J.R. Cuadrado Roura; E. Fuentes Quintana (1990) "El desarrollo económico español y la España desigual de las autonomías". *Papeles de economía española*. 45: 2-61.

Araujo, A; E. Giné (1980) "Topics in Statistical Pattern Recognition". *Communications and Cybernetics* 10. Editado por K. S. Fu. Springer-Verlag.

Ashenfelter, D; Solon, G. (1982) "Longitudinal labor market data-sources, uses and limitations", in: *What's happening to american labor force and productivity of measurements?* Upjohn Institute for Employment Research.

Ashenfelter, O., and J.J. Heckman (1974). "The Estimation of Income and Substitution Effects in a Model of Family Labor Supply". *Econometrica* 42:73-85.

Ashenfelter, Orley; Richard Layard, eds. (1986) *Handbook of Labor Economics*. Amsterdam: North-Holland.

Aurioles Martín, J. (1993) "Andalucía 1986-1992: Balance de situación". *Papeles de economía española*. 55: 84-100.

Aurioles Martín, J.; Martín Rodríguez, M. (1990) "Economía Andaluza 1985-1989. De la crisis a la recuperación". *Papeles de economía española*. 45: 75-98.

Barnett, W.A. (1977). "Pollack and Wachter on the Household Production Function Approach". *Journal of Political Economy* 85:1073-82.

Becker, Gary S. (1965). "A Theory of the Allocation of Time". *Economic Journal* 75:493-517.

Becker, Gary S. (1974). "A Theory of Marriage" in T. W. Schultz, ed., *Economics of the Family*, pag, 293-344. Chicago: University of Chicago Press

Becker, Gary S.(1975) "A theory of Social Interactions". *Journal of Political Economy* 82: 1063-93.

Bennedetti, J (1977) "On the nonparametric estimation of regression function".. *J. Roy. Stst. Soc B* 39:248-253.

Bennet, G. (1962) "Probability inequalities for the sum of independent random variables". *J.A.S.A.* 57: 33-45.

Berkowitz, Monroe; M. Anne Hill, eds. (1986) *Disability and the labor market: Economics problems, policies and programs*. Ithaca, N.Y.: I.L.R. Press.

Berliner, Joseph S. (1989) "Soviet female labor participation: A regional cross-section analysis", *Journal of comparative economics*. 13:446-472.

Berlinet, A. (1979) "Sur les methodes splines en estimation de la densité". *C. R. Acad. Scie. Paris* 288A:847-850.

Bishop, Y.M.M.; S.E. Fienberf; P.W. Holland (1975) "*Discrete Multivariate Analysis: Theory and Practice*". M.I.T. Press. Cambridge.

Blau, Francine; Marianne Ferber (1986) *The economics of women, men and work*. Englewood Cliffs, N.J.: Prentice Hall.

Boneva, L; D. Kendall; I. Stefanov (1971) "Spline transformations: three new diagnostic aids for the statistical data-analyst" *Journal of the Royal Statistical Society B* 33.

Borus, M. E. (1982) "An inventory of longitudinal data sets of interest to economists". *Review of public data use*, 10:113-126.

Bowen, W.G.; T.A. Finegan (1965) "Labor force participation and unemployment", in: Ross, ed. *Employment policy and the labor market*. University of California Press, Berkeley, CA.

Bowen, W.G.; T.A. Finegan (1969). *The Economics of Labor Force Participation*. Princeton, N.J.: Princeton University Press.

Bry, Gerhard (1959) *The average workweek as an economic indicator*. Ocasional paper 69, National Bureau of Economic Research.

Burdett, K.,and D.T. Mortensen (1978). "Labor Supply under Uncertainty" in R.G. Ehrenberg;ed., *Research in Labor Economics*, Vol 2, pp. 109-58.Greenwich, Conn.:JAI Press.

Cacoullos, T. (1965) "Estimation of a probability density function of a density function" *Annals of Math. Stat.* 33:1065-1076.

Cadarso Suárez, M.C. (1990) *Nuevos aportes a la regresión paramétrica y no paramétrica con datos censurados*. Tesis doctoral. Universidade de Santiago de Compostela.

Caneda, A.; M .S. Otero (1994) "El efecto desánimo y adicional en la tasa de participación femenina agrícola gallega". *Estudios de Economía Aplicada*. VIII Reunión Anual de ASEPELT-ESPAÑA.

Carrau, J.M. (1990) "El perfil coyuntural de las comunidades autónomas". *Papeles de economía española*. 45: 225-244.

Cencov, N.N. (1962) "Evaluation of an unknown distribution density from observations" (in Russian) *Doklady. Akad. Neuk. SSSR* 147:45-48.

Cristobal, J.A.; P. Faraldo Roca; W. González Manteiga (1987) "A classe of linear regression parameter stimators constructed by nonparametric estimation" *Ann. Stat* 15.2 603-609.

Cuervo, A.; Vázquez, J.A. (1990) "Asturias: recuperación con ajustes pendientes". *Papeles de economía española*. 45: 115-134.

Chamberlain, G. (1984) "Panel Data" in: Z. Griliches and M. Intriligator, eds., *Handbook of Econometrics*, Amsterdam: North-Holland 2:1247-1318

Danzinger, S., R. Haveman, and R. Plotnick (1981). "How Income Transfers Affect Work, Savings, and the Distribution of Income". *Journal of Economic Literature* 19: 975-1028.

De Lara Peñaste, C.M. (1993) "La economía canaria: situación y perspectivas". *Papeles de economía española*. 55:137-152.

De Miguel, C.; J.R. Lorente (1984) "La evolución de la población activa en los países de la O.C.D.E. tras el inicio de la crisis económica" *Datos demográficos y prospección de tasas de actividad en España*. Publicada por la Asociación Cultural Hispano-Norteamericana en: Profesiones con futuro. Coloquios, El Escorial, Madrid. pp. 53-81.

Devroye, L.P. (1976) *Nonparametric discrimination and density estimation*. Tesis doctoral. Depto. de electrónica. Universidad de Texas.

Dielman, T.E. (1989) *Pooled Cross-Sectional and time series data analysis*. New York: Marcel Dekker.

Dormont, B. (1989) *Introduction à L'économetrie des données de panel*. Paris: Éditions du C.N.R.S.

Douglas, Paul H. (1934) *The theory of wages*. New York: MacMillan.

Drygas, H. (1976) "Weak and strong consistency of the least squares estimators in multiple regression" *Proc. Nat. Acad. Sci. U.S.A.* 75:3034-3036.

Durand, John D. (1948) *The labor force in the United States 1890-1960*. New York: Social Sciences Research Council.

Elliot, Robert F. (1991) *Labor Economics. A comparative text*. McGraw-Hill.

Engel, Robert F.; C.W.J. Granger (1987) "Co-integration and Error Correction: Representation, Estimation and Testing". *Econometrica* 55: 251-76.

Estévez, X (1985) *El Nacionalismo Vasco en el Concierto de los Nacionalismos Peninsulares*. Euskal Herria. Historia eta Gizartea. Historia y sociedad. San Sebastián.

Faraldo Roca P.; W. González Manteiga (1987) "On efficiency of a new class of linear regression estimators obtained by preliminary non-parametric estimation" in: *New perspectives in theoretical and applied statistics*. Puri, Vilaplana, Wertz, eds., John Wiley.

Faraldo Roca, P. (1982) "Una generalización del método de mínimos cuadrados utilizando procedimientos no paramétricos" *IX Jornadas Matemáticas Hispano-Lusas*. Salamanca.

Faraldo Roca, P. (1984) *Nuevos aportes de métodos no paramétricos a la teoría de la regresión paramétrica*. Tesis doctoral. Universidade de Santiago de Compostela.

Fernández-Jardón, C.M.; M.S. Otero. (1993) "El efecto del nivel educativo en la tasa de participación femenina en España". *Estudios de Economía Aplicada*. VII Reunión Anual de ASEPELT-ESPAÑA.

Finegan, T. A. (1962) "Hours of work in the United States - a cross-sectional analysis", *Journal of Political Economy*, 70 (5):452-470.

Finegan, T.A. (1981). "Discouraged Workers and Economic Fluctuations". *Industrial and Labor Relations Review* 35:88-102.

Fisher, R.A. (1922) "On the mathematical foundations of theoretical-statistics" *Philosophical transactions of the Royal Society of London. Series A* 222:309-368.

Fusi, Juan Pablo; (ed.) (1985) *España*. Tomo V: "Autonomías". Madrid. Espasa Calpe.

García Delgado, J.L. (1990) "Madrid: una expansión menos diferente y más condicionada". *Papeles de economía española*. 45: 285-302.

Gasser, T.; H.G. Muller (1979) "Kernel estimation of regression function" *Smoothing techniques for curve estimation L.N.M.* 757, Gasser and Rossemblat, eds, Springer. 23-68.

Gauss, C.F. (1823) "German reference reprinted in Werke English translation "by H. Trotter, *statistical techniques research group*. Princeton University 82-126.

Georgiev, A. (1983) "Local properties of function fitting estimates with application to system identification" *Proc. of the fourth Pannonian Symo. on Math. Stat Bäd Tatzmansdorf*. Austria. Grossman, Plflug, Vineze and Wertz, eds., 141-151.

Georgiev, A. (1985) "Propietes asymtotiques d'un estimateur fonctional non paramétrique" *C.R. Acad. Sc. Paris*, t. 300 serie I, 12:407-410.

González Manteiga, W. (1982) *Construcción axiomática, consistencias y distribuciones de estimadores no paramétricos para funciones de densidad y de regresión*. Tesis doctoral. Universidade de Santiago de Compostela.

González Manteiga, W. (1988) "Una perspectiva general con nuevos resultados de la aplicación de la estimación no paramétrica a la regresión lineal". *Estadística Española*. 30 (118):141-179.

González Manteiga, W. (1990) "Aplicación de la suavización no paramétrica del tipo "K-puntos próximos" a modelos de regresión lineal". *Trabajos de estadística*. 5 (1): 53-67.

González Manteiga, W.; J.M. Vilar Fernández (1987a) "Una clase de estimadores para los parámetros de un proceso AR(1), obtenidos a partir de estimaciones no paramétricas previas" *Trabajos de Estadística* 2 (2): 55-70.

González Manteiga, W.; J.M. Vilar Fernández (1987b) "A class a non-parametrically constructed parameter estimators for a stationary autorregresive model". *Proceedings of the sixth Pannonian Symp. on Math. Stat. Bad Tatzmansdorf*. Reves and Wertz, eds., Reidel-Holland.

Granger, C.W.J. (1983) "Co-integrated Variables and Error-Correcting Models" Unpublished University of California San Diego Discussion Paper 83-13.

Graunt, John; (1662) "Natural and Political Observations on the Bill of Mortality"

Härdle, W. (1990) *Applied Nonparametric Functional estimation*. Cambridge: Cambridge University Press.

Härdle, W.; P. Hall; J.S. Marron (1988) "How far are automatically chosen regression smoothing parameters from their optimun?" *Journal of the American Statistical Asociation* 83: 86-95

Hartog, J.; G. Ridder; J. Theeuwes, eds. (1990) *Panel Data and Labor Market Studies*. Amsterdam: North-Holland.

Hausman, Jerry A. (1978) "Specification test in econometrics", *Econometrica*, 46 (6):1251-1272

Heckman, James J.; Mark Killingsworth; Thomas E. Macurdy (1981) "Empirical evidence on static labour supply models: a survey of recent developments", in: Z.

Horstein, J. Grice and A. Webb, eds, *The economics of the labour markets*. London: Her Majesty's Stationery Office, 74-122.

Heckman, James J.; Singer, B. (1985) *Longitudinal analysis of the labor market data*. Cambridge: Cambridge University Press.

Heckman, James J.; Thomas E. Macurdy (1981) "New methods for estimating labor supply functions: a survey", *Research in labor economics*, 4:65-102.

Hicks, J.R. (1932). *The Theory of Wages*. London: MacMillan Press.

Hoerl, A.E.; R.W. Kennard (1970) "Ridge regression: Biased estimation for non orthogonal problems". *Technometrics*. 12: 55-67.

Hsiao, Cheng (1985) "Benefits and limitations of panel data". *Econometrics Review* 4:121-174.

Hsiao, Cheng (1986) *Analysis of panel data*. Cambridge: Cambridge University Press.

Jevons, W. Stanley (1888) *The theory of political economy*. London: Macmillan.

Jones, Ethel B. (1961) "Hours of week in the United States 1900-1957", *unpublished Ph. D. dissertation*, University of Chicago.

Judge, George J, et al. (1985) *The theory and practice of econometrics*, second edition. New York: John Wiley and Sons.

Kauffman, Bruce E. (1991) *The economics of labor markets*. Orlando: The Dryden Press.

Keeley, Michael C. (1981) *Labour supply and public policy: a critical review*. New York: Academic Press.

Killingsworth, Mark R. (1981) "A survey of labor supply models: theoretical analysis and first-generation empirical results", *Research in Labor Economics*, 4: 1-64.

Killingsworth, Mark R. (1983) *Labor Supply*. Cambridge: Cambridge University Press.

Killingsworth, Mark, R.; J.J. Heckman (1986) "Female labor supply" in: O. Ashenfelter; R. Layard, eds., *Handbook of labor Economics*. Amsterdam: North-Holland. 103-204.

Klevmarken, N.A. (1989) "Panel studies: what can we learn with them? Introduction" *European Economics Review* 33: 523-529.

Kmenta, Jan (1986) *Elements of econometrics*. Second edition. New York: MacMillan.

Kosters, Marvin H. (1966) "Income and substitution effects in a family labor supply model", P-3339, The Rand Corporation.

Lai, T.L.; H. Robbins; C.Z. Wei (1978) "Strong consistency of least squares estimators in multiple regression". *Proc. Nat. Acad. Sci. U.S.A.* 75:3034-3036.

Lancaster, K.J. (1966). "A New Approach to Consumer Theory". *Journal of Political Economy* 74:132-57.

Lazear, Edward (1986) "Retirement from the labor force", in: O. Ashenfelter; R. Layard, eds., *Handbook of labor Economics*. Amsterdam: North-Holland. 1:305-356

Legendre, A.M. (1805) Ver Haster's Survey (1974)

Lewis, H. Gregg (1957) "Hours of work and hours of leisure", in Proceedings of the Ninth Annual Meeting. *Industrial Relations Research Asociation*, 196-206.

Lofstgaarden, D.O.; C.P. Quesemberry (1965) "A nonparametric estimate of a multivariate density function" *Ann. of Math. Stat.* 36:1049-1051.

Long, Clarence D. (1958) The labor force under changing income and employment. Princeton University Press, General Series N 65. Minner, Jacob (1962) "Labor force participation of married woman: a study of labor supply" in: *Aspects of labor economics. National Bureau of Economical Research*, Princeton University Press, 63-105.

Lundberg, S. (1981). "The Added Worker Effect: A Reappraisal". Working Paper N° 706 National Bureau of Economic Research.

Maddala, G.S. (1971) "The use of variance components models in pooling cross section and time series data" *Econometrica* 39: 341-358

Maddala, G.S. (1987 a) "Recent developments in the econometrics of panel data analysis". *Transportation Research* 21: 302-326.

Maddala, G.S. (1987 b) "Limited dependent variable models using panel data" *The Journal of Human Resources* 22: 307-338.

Maddala, G.S. (ed.) (1992) *The econometrics of panel data*. Aldershot, Eng.: Edward Elgar Publishing Limited.

Markov, A.A. (1900) "*Calculus of probabilities*" (in Russian) St. Petersburg. Ver Harter's Surrey (1974).

Mazzodier, P., ed. (1978) *The econometrics of panel data*. Annales de l'INSEE 30/31.

McElroy M.B. and M.J. Horney (1981). "Nash-Bargained Household Decisions: towards a Generalization of the Theory of Demand " *International Economic Review* 22:333-49.

Meghir et al. (1989) "Female participation and male unemployment duration in Greece" *European Economic Review*. 33: 395-406.

Mincer, Jacob (1962) "Labor force participation of married women: a study of labor supply", in: *Aspects of labor economics*. National Bureau of Economical Research, Princeton University Press, 63-105.

Mincer, Jacob (1963). "Market Prices, Opportunity Costs and Income Effects" in C.F. Christ, M. Friedman, L.A. Goodman, Z. Griliches, A.C. Harberger, N. Liviatan, J. Mincer, M. Nerlove, D. Patinkin, L.G. Telrer, and H. Theil eds.; *Measurement in Economics*, pp 67-82. Stanford, Calif.: Stanford University Press.

Mincer, Jacob (1966). "Labor Force Participation and Unemployment: A Review of Recent Evidence "in R.A. Gordon and M.S. Gordon eds.; *Prosperity and Unemployment* pp:73-112. New York: Wiley.

Montricher, G.M.; R.A. Tapia; J.R. Thompson (1975) "Nonparametric maximum likelihood estimation of probability densities by penalty function methods". *Ann. of Stat.* 3: 1329-1348.

Moore, D.S.; J.W. Jackel (1976) "Large sample properties of nearest neighbor density function estimators". *Statistical decision theory and related topics I*. New York: Academic Press.

Muth, R.F. (1966). "Household Production and Consumer Demand Functions". *Econometrica* 34:699-708.

Novales, A. (1988) "La incorporación de la mujer al mercado de trabajo en España: participación y ocupación". *Moneda y crédito*. 188: 243-281.

Novales, A. (1993) *Econometría*. Madrid: McGraw Hill.

Pampillón, R.; Vega, J.; Espárrago, F. (1990) "La economía extremeña 1985-1989: estructura y problemas". *Papeles de economía española*. 45: 245-262.

Parzen, E. (1962) "On estimation of a probability density function and mode" *Ann. of Math. Stat.* 33:1065-1076.

Pedreño, A. (1988) "Un eje de expansión económica: Cataluña-Mediterráneo". *España-Economía*. Madrid: Espasa-Calpe. Pg: 797 y siguientes.

Pencavel, John (1986) "Labor supply of men: a survey", in: O. Ashenfelter; R. Layard, eds., *Handbook of labor Economics*. Amsterdam: North-Holland, 1:3-102

Pissarides, C.A. (1976). *Labour Market Adjustment*. Cambridge: Cambridge University Press.

Pollack, R.A. and T.J. Wachter (1974). "The Relevance of the Household Production Function and Its Implication for the Allocation of the Time". *Journal of Political Economy* 83:255-77.

Prakasa Rao, B.L.S. (1983) *Nonparametric functional estimation*. New York: Academic Press.

Priestley, M.B.; M.T. Chao (1972) "Nonparametric function fitting" *J. Roy. Stat. Soc. B* 34: 385-392.

Quintás, J.R. (1990) "La economía gallega en el final de los años ochenta". *Papeles de economía española*. 45: 263-284.

Quintela del Río, A. (1992) *Cálculo del parámetro de suavización en la estimación de curvas con datos dependientes*. Tesis doctoral. Universidade de Santiago de Compostela.

Raj, B.; B.H. Baltagi, eds. (1992) *Panel Data Analysis*. Heidelberg: Physica-Verlag.

Raymond, J.L. (1990) "El perfil coyuntural de las comunidades autónomas". *Papeles de economía española*. 45: 62-74.

Reynolds, Lloyd G.; Stanley H. Masters; Coletta H. Mosser (1991) *Labor Economics and Labor Relations*. Englewood Cliffs, N.J.: Prentice Hall.

Robinson, P.M. (1988) "Semiparametric Econometrics: A Survey". *Journal of Applied Econometrics* 3: 35-51.

Rossemblat, M. (1956) "Remarks on some nonparametrics estimates of a density function." *Annals of Math. Stat.* 27: 832-835.

Ruppert, D.; R.J. Carrol (1980) "Trimmed Least Squares Estimation in the Linear Model" *J.A.S.A.* 75: 828-838.

Samuelson, P.A. (1956). " Social Indifference Curves". *Quartely Journal of Economics* 70: 1-22.

Schoemberg, E.; P. Douglas (1937) "Studies in the supply curve of labor: the relation in 1929 between average earnings in american cities and the proportion seeking employment", *Journal of Political Economy*, 45 (1): 45-79.

Serrano Sanz, J.M. (1990) "Aragón: una década de crecimiento". *Papeles de economía española*. 45: 101-114.

Silverman, B.W. (1986) *Density estimation for statistics and data analysis*. London: Chapman and Hall.

Solé Tura, Jordi (1985) *Nacionalidades y nacionalismos en España. Autonomías, Federalismo, Autodeterminación*. Madrid.

Stute, W. (1984) "Asymptotic normality of nearest neighbor regression function estimates" *Ann. Stat.* 12: 917-926.

Stute, W; W. González Manteiga (1990) "Nearest neighbor smoothing in linear regression" *Journal of Multivariate Analysis*. 34: 61-74.

Tachibanaki, T.; K. Sakurai (1991) "Labor suply and Unemployment in Japan" *European Economic Review*. 35: 1575-1587.

Toharia, Luis, ed. (1983) *El mercado de trabajo: Teoría y aplicaciones*. Madrid: Alianza Editorial.

Velasco Barroetabeña, R.; Díez López, M.A.; García Arteche, I. (1990) "Claroscuros en la recuperación y perspectivas de la economía vasca". *Papeles de economía española*. 45: 349-372.

Vilar Fernández, J.M. (1987) *Estimación no paramétrica de la función de densidad y predicción en series de tiempo*. Tesis doctoral. Universidade de Santiago de Compostela.

Villaverde Castro, J. (1990) "La economía de Cantabria: evolución y perspectivas". *Papeles de economía española*. 45: 165-134.

Wahba, G. (1971) "A polynomial algorithm for density estimation" *Ann. of Math. Stat.* 42: 1870-1886.

Wahba, G. (1975) "Interpolating spline methods for density estimation" I. Equi-spaced Knots. *Ann. Stat.* 3:30-48.

Willis, Robert J. (1987) "What have we learned from the economics of the family", *American Economic Review* 77 (May 1987): 68-81.

Woytinsky, W.S. (1940) *Additional workers and the volume of unemployment in the depression*. New York: Social Sciences Research Council.

UNIVERSIDADE DA CORUÑA
Servicio de Bibliotecas



1700757467